DOCUMENT RESUME

ED 348 570 CE 061 829

AUTHOR Hoachlander, E. Gareth; And Others
TITLE Accountability for Vocational Education: A Practitioner's Guide.
INSTITUTION National Center for Research in Vocational Education, Berkeley, CA.
SPONS AGENCY Office of Vocational and Adult Education (ED), Washington, DC.
PUB DATE Jul 92
CONTRACT V051A80004-92A
NOTE 133p.; One in a series of Technical Assistance Reports.
AVAILABLE FROM NCRVE Materials Distribution Service, Horrabin Hall 46, Western Illinois University, Macomb, IL 61455 (order no. MDS-407: $10.25).
PUB TYPE Guides - Non-Classroom Use (055)

EDRS PRICE MF01/PC06 Plus Postage.
DESCRIPTORS *Accountability; Behavioral Objectives; Educational Policy; Evaluation Criteria; Performance; *Performance Factors; Postsecondary Education; Secondary Education; Special Needs Students; *Standards; *Student Evaluation; *Vocational Education

ABSTRACT

This guidebook offers guidelines, suggestions, examples, and recommendations for developing a comprehensive, sound system of accountability. The first chapter develops a conceptual framework for developing systems of accountability that identifies some major policy issues that must be addressed. Chapter 2 discusses getting started and identifies some basic decisions that must be made at the outset. Chapter 3 examines specific examples of performance measures and assesses the pros and cons of different types of measures. It also discusses appropriate distinctions between secondary and postsecondary vocational education. Chapter 4 turns to a discussion of performance standards, including examples for different standards and procedures for determining them. Chapter 5 describes different approaches for introducing incentives and adjustments for special populations. Chapter 6 considers issues of implementation and includes possibilities for phasing in accountability systems and monitoring and modifying them over time. Chapter 7 provides information on student assessment in the context of performance measures and standards. A glossary is included as an appendix. (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED348570

NCRVE

National Center for Research in
Vocational Education

University of California, Berkeley

Accountability for Vocational Education: A Practitioner's Guide

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)

☐ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

CE 061829

Supported by
the Office of Vocational and Adult Education,
U.S. Department of Education

BEST COPY AVAILABLE

ACCOUNTABILITY FOR VOCATIONAL EDUCATION: A PRACTITIONER'S GUIDE

E. Gareth Hoachlander
Karen Levesque
Mikala L. Rahn

University of California at Berkeley

National Center for Research in Vocational Education
University of California at Berkeley
1995 University Avenue, Suite 375
Berkeley, CA 94704

Supported by
The Office of Vocational and Adult Education,
U.S. Department of Education

July, 1992 MDS-407

This document is one of a series of Technical Assistance Reports. This document has not been reviewed by NCRVE; therefore, this paper represents the views of its authors and not necessarily those of the Center or the U.S. Department of Education. NCRVE makes Technical Assistance Reports available, upon request, for informational purposes.

FUNDING INFORMATION

Project Title:	National Center for Research in Vocational Education
Grant Number:	V051A80004-92A
Act under which Funds Administered:	Carl D. Perkins Vocational Education Act P. L. 98-524
Source of Grant:	Office of Vocational and Adult Education U.S. Department of Education Washington, DC 20202
Grantee:	The Regents of the University of California National Center for Research in Vocational Education 1995 University Avenue, Suite 375 Berkeley, CA 94704
Director:	Charles S. Benson
Percent of Total Grant Financed by Federal Money:	100%
Dollar Amount of Federal Funds for Grant:	$5,775,376

Disclaimer:

This publication was prepared pursuant to a grant with the Office of Vocational and Adult Education, U.S. Department of Education. Grantees undertaking such projects under government sponsorship are encouraged to express freely their judgement in professional and technical matters. Points of view of opinions do not, therefore, necessarily represent official U.S. Department of Education position or policy.

Discrimination:

Title VI of the Civil Rights Act of 1964 states: "No person in the United States shall, on the ground of race, color, or national origin, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any program or activity receiving federal financial assistance." Title IX of the Education Amendments of 1972 states: "No person in the United States shall, on the basis of sex, be excluded from participation in, be denied the benefits of, or be subjected to discrimination under any education program or activity receiving federal financial assistance." Therefore, the National Center for Research in Vocational Education project, like every program or activity receiving financial assistance from the U.S. Department of Education, must be operated in compliance with these laws.

ACCOUNTABILITY FOR VOCATIONAL EDUCATION: A PRACTITIONER'S GUIDE

TABLE OF CONTENTS

Chapter 1: Conceptual Framework 1

Chapter 2: Getting Started 21

Chapter 3: Performance Measures 45

Chapter 4: Performance Standards 79

Chapter 5: Incentives and Adjustments for Special Populations 91

Chapter 6: Implementation Issues 107

Chapter 7: Student Assessment 115

Glossary 127

NCRVE

National Center for Research in Vocational Education
1995 University Avenue, Suite 375
Berkeley, California 94704-1058
(510) 642-4004

ACKNOWLEDGMENTS

The National Center for Research in Vocational Education (NCRVE) convened a Technical Advisory Group to advise on upcoming technical assistance offered to states on developing performance measures and standards. The group met November 6–8, 1991 to discuss technical assistance activities. In the following months, NCRVE staff developed a conceptual framework and guidelines on developing performance measures and standards that were reviewed by the Technical Advisory Group at its second meeting January 29–31, 1992. The feedback from this meeting enabled NCRVE to develop a draft of *Accountability for Vocational Education: A Practitioner's Guide* that was distributed to the participants in three regional workshops held in March 1992.

We are very appreciative of the essential input and time the members contributed. Without the efforts of the following individuals, NCRVE would not have attained its current depth of knowledge on performance measures and standards.

Sri Ananda, Far West Laboratory for Education Research and Development
Don Brannon, North Carolina State Department of Education
Joe Casello, Office of Vocational and Adult Education, U.S. Department of Education
Norm Constantine, Far West Laboratory for Education Research and Development
Evelyn Ganzglass, National Governors Association
John Glenn, Iowa State Board of Education
Don Hatton, Electronic Industries Association
Charles Hopkins, Oklahoma State Department of Education
Lauren Jacobs, Center for Law and Education
Mel Johnson, Minnesota Technical College Board
John Klit, Illinois State Department of Education
Charles Losh, Arizona State Department of Education
Kathy Nicholson-Tosh, Illinois State Department of Education
Debra Nolan, Office of Vocational and Adult Education, U.S. Department of Education
Roy Peters, Oklahoma State Department of Education
Barbara Shay, New York State Department of Education
Patricia Stanley, Chancellor's Office, California Community Colleges
Brian Stecher, RAND Corporation
Walter J. Seymour, Electronic Industries Association
Twila Young-Glenn, South-Central Federation of Labor, AFL-CIO

NCRVE members: Gary Hoachlander, Phyllis Herriage, Carolyn Maddy-Bernstein, Karen Levesque, and Mikala Rahn.

In addition, we would like to thank the participants in the regional workshops in March 1992. The questions and information shared by participants were essential to the completion of this guide.

Many individuals made substantial contributions to the production of this guide. Andrea Livingston, Brett Threlkeld, and David Donald edited the guide. Leslie Retallick and Benjamin Pinney preparedand formatted the text and graphics. NCRVE's Materials Distribution Servicefurnished invaluable assistance in editing, production and shipping. Special thanks are due Pam Mainland, Sandra Larimer, and Diana Burnell.

1. Conceptual Framework

The Carl D. Perkins Vocational and Applied Technology Education Act of 1990 requires states to develop accountability systems that include performance measures and standards for secondary and postsecondary vocational education programs. These systems are to include at least two sets of performance measures. One set must be measures of learning and competency gains, including student progress in the achievement of basic and more advanced academic skills. The other set must include any one of the following four measures: (1) competency attainment, (2) job or work skill attainment, (3) retention in school, or (4) placement in further education, the military, or employment. These systems must also include appropriate adjustments and incentives for encouraging services to students with special needs. Finally, states and local eligible recipients may supplement these minimal requirements with additional measures, and local recipients may also modify measures and standards to reflect local demographic or economic conditions.

In some respects, the federal requirements for accountability are quite modest. Only two sets of measures are required, and states enjoy considerable discretion in identifying and defining measures, determining standards, and incorporating these measures and standards into annual program evaluations. Moreover, the law gives states two years to develop their accountability systems, calling for implementation by Fall 1992.

On the other hand, a careful reading of the Act and a thoughtful interpretation of its many provisions can complicate even the law's minimum requirements. For example, the law encourages vocational education to provide students with "strong experience in and understanding of all aspects of the industry the students are preparing to enter."[1] Thus, it seeks to significantly broaden the kinds of knowledge and skills students acquire through participation in vocational education. If assessing students' acquisition of more conventional academic skills and occupational competencies is relatively straightforward (and in some instances, it decidedly is not), expanding evaluation to include attention to this broadened mission of vocational education will pose a number of difficult problems.

A careful reading of the Act and a thoughtful interpretation of its many provisions can complicate even the law's minimum requirements.

Most states want to go well beyond the minimum requirements of Perkins (Hoachlander and Rahn, 1992).[2] They recognize that limiting accountability

[1] The law defines "all aspects of the industry" as meaning "strong experience in, and understanding of, all aspects of industry students are preparing to enter, including planning, management, finances, technical and production skills, underlying principles of technology, labor and community issues, and health and safety." (Sec. 521 (2)).

[2] In the National Center for Research in Vocational Education's (NCRVE) survey of the states in Summer 1991, only two states indicated that they would limit their systems to the two types of measures required by Perkins, thirty states said they would be developing four or more measures, and fifteen planned to use seven or more.

to just two measures and standards would unwisely displace a number of other important goals of vocational education. Whatever one of the four optional goals states might elect in order to comply with the minimum requirements of Perkins, the other three would remain important. Yet in excluding them from the accountability systems, states would send an implicit message that these other outcomes no longer mattered or that their importance had considerably diminished. Consequently, many states are planning to incorporate all five of the measures outlined by Perkins—and in some cases, even more measures—in their accountability systems. Developing a comprehensive, sound system of accountability, therefore, poses a major challenge.

In the chapters that follow, we offer some guidelines, suggestions, examples, and recommendations for meeting this challenge. While the law prescribes the basic parameters that states must follow in designing accountability systems and conducting local evaluations, it provides states with considerable discretion to develop systems that will fit their own particular situations or needs. The law is quite permissive, especially when a good rationale underlies the actions a state or locality wants to take.

While the law prescribes the basic parameters that states must follow, it provides states with considerable discretion to develop systems that will fit their own particular situations or needs.

Nevertheless, there is little point in reinventing the proverbial wheel, and the ideas presented here should save states and local recipients some time. While we are not advocates of comparability and uniformity for their own sake, common definitions and procedures will make it easier to understand outcomes from a state, regional, or national perspective. We believe that measures and standards should first serve local needs, but if certain uniform measures and standards can satisfy local requirements, then uniform and comparable definitions are preferable to those that are not.

This, then, is a guidebook, not a litany. There is no one right approach to designing systems of accountability. There are many choices that involve difficult tradeoffs—between accuracy and timeliness, precision and consistency, simplicity and fairness, breadth and burden, knowledge and cost. States and localities will elect to make these choices in different ways and for different reasons. As long as they make these choices with one overriding objective in mind—program improvement—they are not likely to go wrong.

The guidebook is organized into seven chapters. This first chapter develops a conceptual framework for developing systems of accountability that identifies some of the major policy issues that must be addressed. Chapter 2 discusses getting started and identifies some of the basic decisions that must be made at the outset. Chapter 3 examines specific examples of performance measures and assesses the pros and cons of different types of measures. It also discusses appropriate distinctions between secondary and postsecondary vocational education. Chapter 4 turns to a discussion of performance standards, including examples for different standards and procedures for determining them. Chapter 5 describes different approaches for introducing incentives and adjustments for special populations. Chapter 6 considers issues of implementation and includes possibilities for phasing in accountability systems and monitoring and modifying them over time. Chapter 7 provides information on student assessment. A glossary is included as an appendix.

The New Accountability

The emphasis of the current legislation on accountability is not entirely new to federal vocational education policy. Beginning with the Vocational Education Act of 1963, states were required to develop state plans and conduct program evaluations. Amendments in 1968, 1976, and 1984 added to these planning requirements. Indeed, the 1976 Amendments explicitly required states to submit an annual Accountability Report. Moreover, earlier legislation paid some attention to student outcomes. For example, federal law has long encouraged using various labor market outcomes, such as placement rates or employer satisfaction, to assess vocational education programs.

What is new about the 1990 Act is a much stronger emphasis on student *outcomes* as the focus for evaluation and planning. Traditionally, planning and evaluation have been primarily concerned with issues of compliance and have attended mainly to procedural aspects of implementing federal vocational education policy. On the other hand, a long-standing criticism of the federally required State Plans for Vocational Education is that these plans amount to little more than compliance documents that merely parrot back assurances that federal law and regulations will be observed. The new law seeks to redirect attention toward student and program outcomes, more broadly conceived. Thus, not only does the new law encourage more effective use of labor market outcomes in assessing program effectiveness and student progress, but also it requires attention to *learning* outcomes. Learning outcomes include both occupational competencies and measures of academic achievement. The law broadens conceptions of occupational competencies to include attention to more "general occupational skills" that will provide a foundation not only for a specific job but for a lifetime of adapting to major changes in the workplace.

What is new about the 1990 Act is a much stronger emphasis on student outcomes.

This shift from process-based to outcome-based policy presents important opportunities as well as major challenges for secondary and postsecondary vocational education. Perhaps most importantly, emphasis on outcome-based policies invites serious and widespread debate about the purposes of vocational education and its relation to the rest of the educational system. Precisely what does vocational education seek to accomplish and whom does it seek to serve? What kinds of skills does vocational education seek to impart? How specific should these skills be to particular industries or occupations? What is the appropriate mix of academic and occupational skills? How should broader definitions of general occupational skills be conceived, and what constitutes a useful and practical understanding of "all aspects of the industry?" How do the answers to such questions differ for secondary and postsecondary education? Any serious consideration of outcome measures and performance standards must grapple with these weighty questions.

Federal policy offers only the most general directions for answering these questions and leaves to the states the task of designing accountability systems. Hence, there will be at least fifty different approaches to system design, with even more permutations to accommodate differences between

secondary and postsecondary programs or among local eligible recipients. Decentralized planning and implementation are not inherently bad. However, in the absence of some strong guiding principles and attention to potentially common state and local interests, there is a high probability of chaos that could discredit and undermine basically sound concepts.

It should be stressed that what the federal government seeks from better accountability, planning, and evaluation is not always clear, and this lack of clarity adds to confusion at the state and local level. The federal government often sends mixed messages. Thus, the primary purpose of better accountability is said to be local program improvement, which tends to argue for more decentralized systems tailored to local needs and conditions. On the other hand, there are national pressures for world class standards and national exams that appear to call for more uniform systems with comparable definitions. These two aims can have much in common, but they also create tensions that can complicate developmental efforts.

To some extent, the mixed message regarding the purposes of accountability reflects the multiplicity of actors who have an interest in accountability. These actors include students, parents, employers, local educators, state legislators, state administrators, federal officials, members of Congress, and other interested parties. It is probably not possible to design an accountability system that will satisfy all of these parties and to attempt to do so may produce a system that satisfies no one.

The primary purpose of the accountability requirements of Perkins must be local program improvement.

We have assumed that the primary purpose of the accountability requirements of Perkins must be local program improvement—helping local educators to develop the management and information systems they require to make vocational education respond more effectively to the needs of students. This assumption means that the systems that are developed must accommodate the diversity that characterizes vocational education and maximize adaptation to local needs and circumstances. It also means that hierarchical systems that seek to aggregate uniform and comparable data are not likely to work. Better state and national data on vocational education may be desirable, but other means will need to be devised to generate that information. Congress did not intend the accountability systems to bear this burden (Hill, Harvey, and Praskac, forthcoming).

Finally, even if some consensus can be achieved on what types of outcomes are appropriate to measure and what levels of performance should be expected, the means for assessing these outcomes fairly, reliably, and validly may be lacking. To date, far more attention has been paid to assessing academic skills than to testing occupational competencies. Moreover, even given the myriad of tests for assessing academic skills, it is by no means clear that any of these tests measure the right things—the ability of students to effectively apply academic skills in a practical work setting rather than merely the ability to manipulate concepts in the abstract. Consequently, the systems designed for Fall 1992 will require continued development and ongoing revision.

In summary, the new accountability requirements of Perkins can help states and localities to improve secondary and postsecondary vocational offerings. Realizing this aim, however, will require careful attention to system design, as well as patience in overcoming a number of conceptual and methodological obstacles.

Basic Requirements for an Effective Accountability System

Developers of state and local accountability systems need to consider at least six basic system requirements that will help to ensure sound design. As development proceeds, it is useful to periodically assess how well accountability systems address these six principles of design:

1. The desired measures and standards should be defined clearly and precisely.

2. Outcomes should be easily and accurately measured, while minimizing data burden and cost.

3. A manageable number of outcome measures should comprise the accountability system.

4. Standards must have validity, justified in terms of success in either the workplace or further education and training; they must also be fair, avoiding bias by race, gender, or special need.

5. Data for each measure should be collected at appropriate intervals of time.

6. The information generated by the accountability system should be routinely accessible by students, teachers, administrators, parents, employers, board members, and others interested in educational policy and performance.

We will briefly elaborate on each of these.

1. Be Clear and Precise

The first step in designing accountability systems is identifying and defining the kinds of outcomes that the system will monitor. Federal law and regulations offer some general directions, but both fall well short of providing clear, precise definitions. For example, the first measure required by Perkins is a measure of student gains, including gains in academic achievement. Is this a measure of both occupational and academic gains? Are the gains to be measured for individual students, programs, or institutions? Should the gains be measured relative to some absolute standard, or may they be calculated based on relative gains, such as expressed as percentage increases?

Chapter 3 will have much more to say about these kinds of measurement issues and will provide some specific examples. Here we simply emphasize that states will need to devote considerable thought to clarifying outcomes and defining them in ways that can be measured precisely.

2. Seek Measures that are Easily Obtained, Accurate, and Efficient

System designers should first consider the utility of existing information before imposing new data requirements.

The process of identifying and defining appropriate outcome measures must consider availability of data, accuracy of available information, and the cost of maintaining and reporting information. Data are rarely free and developing new assessment instruments and data systems can be costly indeed. System designers should, therefore, first consider the utility of existing information before imposing new data requirements. Schools routinely maintain much information that can be used in an outcome-based information system. For example, most secondary schools have information on academic achievement and many even monitor occupational competencies. Virtually all postsecondary institutions maintain records on degrees and certificates conferred in occupational programs. With one or two exceptions, all states through their unemployment insurance systems maintain wage record data that can provide a cheap and fairly comprehensive source of information for student follow-up.

Existing data are not without problems, and there will undoubtedly be shortcomings and gaps. In the short run, most states will be able to meet the minimum requirements of Perkins without developing elaborate new data sources. Over the longer term, however, attention should be paid to new data, especially better information on learning gains. As better assessment instruments are developed to measure students' academic and occupational skills (including students' understanding of all aspects of the industry), new measures of student achievement should be incorporated into the accountability systems.

3. Select a Manageable Number of Measures

Perkins requires only two sets of measures for accountability—measures of learning gains and one of four other possible options. While it is clear that only two sets of measures will be sufficient to comply with federal law, it is doubtful that only two sets will result in a sound accountability system. Vocational education has multiple objectives. It seeks to impart a wide variety of academic, occupational, and employability skills. It aims to keep students in school and encourages them to complete occupational programs and academic requirements for high school graduation or postsecondary degrees and certificates. Ideally, program completion should lead to gainful employment or successful matriculation in further education or training. Stressing any one of these objectives over another is likely to lead to neglect of other important aims. Hence, it is important to recognize multiple objectives in system design.

How many measures are appropriate? There is no one right answer. In our view, five or six measures are probably the minimum. Anything less and some important aims are likely to be ignored. On the other hand, more than twelve measures are likely to become unmanageable, significantly increasing data burden and quite possibly diffusing the mission of vocational education so broadly that it lacks focus and direction. It should be emphasized that systems do not have to implement every measure at the same time. States can choose to phase in multiple measures over time. Chapter 6 will discuss options for phasing in accountability in detail.

4. Validate Standards

One of the most difficult challenges states face in designing accountability systems is the specification of appropriate standards. Law and regulations are completely silent on this topic. Chapter 4 will examine standard setting at length. Here, we simply stress that whatever standards are adopted, they must have validity in the workplace or in further education or training. Standards that are arbitrary and lack a clear rationale will be suspect and subject to challenge, especially if they have the effect of screening students or withholding the award of degrees and certificates. Hence, in setting standards, states must constantly ask themselves what evidence is there that achievement of this level of competence is necessary to perform successfully in the workplace or to advance to a higher level of education and training?

Whatever standards are adopted must have validity in the workplace or in further education or training.

Closely related to validity is the issue of fairness. It will be necessary to demonstrate that the measures adopted and the assessment instruments used to produce these measures are free of bias.

5. Determine Appropriate Frequency of Collection

Different measures have different implications for the frequency with which data must be collected and reported. Most of the data will probably not need to be collected and analyzed more than once a year. However, particularly at the local level, it may prove desirable to collect data more frequently. For example, suppose a local school district chooses to improve student attendance. Although student attendance is not an outcome measure, attendance at school is certainly a necessary precondition for learning to occur. Moreover, student attendance can be a measure of retention in school. A school desiring to monitor its performance on improving student attendance may want to analyze results monthly or even weekly and might announce weekly attendance at the entrance to the school as a way of keeping students and faculty focused on attendance goals. Similarly, local recipients may opt to monitor student achievement more frequently than once a year.

6. Make Performance Information Readily Available

Accountability will have little importance if information on performance is not made available to those able to use it to make decisions about program planning and program improvement. Clearly, teachers, administrators, and board members need to review the data generated by accountability systems, as do legislators and others interested in educational policy. Equally important, students and parents should have access to information on the performance of various programs and institutions. Students are making important choices about how to invest their time, and these choices should be well informed. Consequently, an important feature of system design is attention to reporting formats, including appropriate safeguards against the misuse or abuse of performance information.

Relationship to Local Evaluation

"Each recipient . . . shall annually evaluate the effectiveness of the program . . . based on the standards and measures . . . developed as required by Section 115."

PERKINS ACT Section 117

The measures and standards developed for the state's accountability systems are to play a major role in local evaluation. Section 117(a) of the Perkins Act states that each recipient of funds under the basic grant "shall annually evaluate the effectiveness of the program conducted with assistance under this Act based on the standards and measures (or modifications thereto) developed as required by section 115." Recipients with programs that are not making "substantial progress" in meeting the standards established by the state must then develop a plan for program improvement that must be developed in consultation with teachers, parents, and students and include a description of the strategies that will be used to improve programs.

In addition to assessing progress toward satisfying the standards established as part of the statewide accountability system, the annual local evaluation must also examine the access and success of members of special populations in vocational education programs. Furthermore, the evaluation must assess the progress of vocational education programs in providing students with strong experience in and understanding of all aspects of the industry the students are preparing to enter.

States may choose to incorporate all or some of these additional features of local evaluation into their accountability systems. For example, states may decide that the measures adopted as part of the system should be reported for various special populations, provided local reporting units have the capability to define and identify members of special populations in a uniform and comparable fashion.[3] Similarly, as states develop operable, quantifiable definitions of knowledge of "all aspects of the industry," measures and standards with respect to this variable could be incorporated directly into accountability systems. Nevertheless, regardless of whether these additional criteria are made part of the statewide accountability system, they must be addressed by recipients of the basic grant as part of their annual evaluations.

[3] This provision can be difficult to satisfy. For example, operable definitions for economically disadvantaged students, as well as practical methods for identifying them, are far from universally established.

It should be noted that the law defines neither "substantial progress" toward meeting standards nor "lower rates" of access by members of special populations. Rather the legislation leaves it to eligible recipients to describe in their local applications the evaluation standards that the applicants will use to measure their progress. Presumably, states also can establish standards for improvement. For example, a state could establish a program improvement standard that would define acceptable progress as "in any given year, diminishing by ten percent or more the gap between the state standard for achievement and the actual achievement of a recipient performing below that state standard."

The legislation leaves it to the eligible recipients to describe in their local applications the evaluation standards that the applicants will use to measure their progress.

Similarly, states could also define criteria for what constitutes a balanced program or program outcome with respect to race, gender, and membership in special populations. Presumably, such criteria would be defined relative to some level of acceptable deviation from statewide, community, or institutional norms. For example, assuming women represent fifty percent of the student population, a balanced program might be defined as one that contains between forty and sixty percent women.

Measures versus Standards

What is a performance *measure* and how does it differ from a performance *standard*? This a frequently asked question by designers of accountability systems.

A performance *measure* is the type of outcome that is considered appropriate for monitoring. Measures such as proficiency in reading, writing, math, and science—as exhibited by scores on standardized achievement tests—constitute measures of academic achievement. A score on a competency-based test is an example of a measure of occupational competency. The percentage of students finishing a designated sequence of academic and occupational studies is a measure of program completion. The percentages of program completers placed in jobs, further education, or the military are all measures of placement.

A performance *standard* is a level of performance that individual students, programs, or institutions are expected to achieve. For example, "students will be expected to score eighty percent or higher on tests of reading, writing, and mathematics in order to successfully complete the automotive technology program" is a standard for determining students' readiness to work or to continue to a more advanced level of education or training. Similarly, "programs will maintain placement rates of seventy percent or higher" is a standard states or locals might use to determine whether or not a program should be maintained.

Types of Measures

As states develop their accountability systems, there are at least three general types of measures they will want to consider: (1) learning measures,

(2) labor market outcomes, and (3) measures of access and participation. These will be discussed briefly here, but with more detail provided in the chapters that follow.

Learning Measures

"Each system . . . shall include . . . measures of learning and competency gains, including student progress in the achievement of basic and more advanced academic skills."

PERKINS ACT
Section 115

The law requires that states develop measures of learning and competency gains and explicitly states that these must include measures of student progress in the achievement of basic and more advanced academic skills. Implicit in this language is the requirement that measures of occupational skills must also be included, although the law does not clearly say so. Nevertheless, given that a primary purpose of vocational education is preparing students for employment, measures of occupational competency are clearly a desirable component of the accountability system.[4]

The key word in the law's language is *gains*, evidence that participation in vocational education leads to an increase in knowledge and skills. There are at least two approaches to measuring gains. First, some form of pretesting and posttesting can be used. This method is student-based and assesses changes for individual students in knowledge and skills over some specified period of time—between Fall and Spring or eighth grade and eleventh grade, for example. Second, gains can be measured programmatically over time. Such an approach averages scores of students participating in a program at one point in time and compares this average with an average from an earlier period to determine whether scores have improved. This programmatic approach requires statistical controls for changes in student characteristics and in other variables external to the program that may influence achievement.

In contrast to gains, one can also measure learning *attainment*, the actual level of competency (academic or occupational) that a student has mastered. Competency attainment and job or work skill attainment are two of the four options states have in developing their second set of required performance measures.

In summary, then, learning measures assess gains or attainment with respect to academic and occupational knowledge and skills. Some examples of possible learning measures include:

- Student achievement on standardized tests
- Student achievement on competency-based tests

[4] As desirable as measures of occupational competency may be, many states may encounter serious difficulty in trying to develop appropriate measures. Occupational competency testing is not as well developed, nor as widely standardized, as academic achievement testing. Moreover, there are as yet no good assessment instruments for assessing general occupational skills that represent students' understanding of "all aspects of the industry." Consequently, developing good measures of occupational competency is likely to prove difficult for the majority of states that do not presently conduct statewide occupational testing.

- Course completion rates
- Program completion rates
- High school graduation rates
- Rates of completing degrees or certificates
- Patterns of course taking—for example, the percentage of students pursuing a planned sequence of vocational and academic courses as evidenced by student transcripts.

Labor Market Outcomes

Measures of labor market outcomes are familiar to most vocational educators, having been used for some time to evaluate program effectiveness. Among those measures that have been traditionally employed are the following:

- Time needed to find and secure employment
- Placement rates
  - In a job related to training
  - In any job
  - In further education or training
  - In the military or other service (e.g., the Peace Corps)
- Unemployment rates
- Entry-level wages
- Quarterly earnings
- Time employed in first job
- Rate at which quarterly earnings increase
- Employer and employee satisfaction.

Access Measures

Access by members of special populations to high quality vocational education is a long-standing concern of federal policy. Traditionally, access has been examined primarily in terms of enrollment in vocational education programs. While program *participation* continues to be an important indicator of access, program *outcomes* are assuming equal importance. Hence, how well members of special populations fare with respect to learning gains and labor market success adds a new dimension to conceptions of access. States that have the ability to track the performance of individual students and identify those with special needs will find it useful to incorporate

How well members of special populations fare with respect to learning gains and labor market success adds a new dimension to conceptions of access.

measures of access into their accountability systems. Some examples of possible access measures include the following:

- Ratio of percentage of students with special needs *enrolled* in selected vocational education programs to the percentage of students with special needs in the school population at large;
- Ratio of percentage of students with special needs *completing* selected vocational education programs to the percentage of students without special needs completing selected vocational education programs;
- Ratio of percentage of students with special needs *entering jobs* related to training to the percentage of students without special needs entering jobs related to training;
- Ratio of average achievement test scores by students with special needs enrolled in vocational education programs to average of students with special needs not participating in vocational education.

All of these measures can be calculated and reported for each type of special population, including gender and race–ethnicity.

Cross-Sectional Versus Longitudinal Measures

There are two basic approaches, cross-sectional and longitudinal, for collecting data on performance measures. A cross-sectional approach collects information across a specified population at a single point in time. Tenth grade achievement test scores are an example of cross-sectional data collection. In contrast, a longitudinal approach collects information on the same individuals at different points in time. For example, a system that maintains data for each student on test scores in the eighth, tenth, and twelfth grades is longitudinal and allows the development of longitudinal measures of changes over time.

Longitudinal systems require the ability to collect, store, and analyze data on individual students.

Generally, longitudinal data systems are more complex and more costly than cross-sectional systems, requiring the ability to collect, store, and analyze data on individual students. Cross-sectional systems, in contrast, typically aggregate student level data to a higher level—for example, the classroom or a single institution. Consequently, data storage and analysis are easier, and while records can be maintained over time, there is no inherent requirement for doing so.

While more costly and complex, well-designed longitudinal data systems are a richer and more useful source of data for program improvement. They facilitate direct measures of student gains and can generally contribute to more sophisticated analyses of program effectiveness.

Attainment Versus Gains

As noted above, federal law requires states to adopt measures of gains as part of their accountability systems. While not required, attainment measures are some of the four options states may elect as the second type of measure. In our view, it is important that measures of gains and attainment both be included in systems of measures and standards. Measures of gains, or "value added," will help to discourage creaming and other forms of screening out students less likely to succeed immediately. In the absence of measures of value added or changes over time, educators may be encouraged to admit only the best students to courses and programs; this temptation becomes greater the more heterogeneous the eligible participants are in terms of ability and previous preparation. Measures of change over time also provide evidence of the degree of improvement being achieved by different programs and institutions. Finally, gains measures ensure that initially low-performing programs or institutions have the opportunity to demonstrate progress over time, instead of being held to an absolute standard that may not be attainable in a short period.

It is important that measures of gains and attainment both be included in systems of measures and standards.

Nevertheless, attainment is important. If particular levels of attainment are necessary to succeed in the labor force, performance that is anything short of those levels will not lead to labor market success—no matter how great the gains over time. Educators need to clearly communicate to students the level of performance they must attain for success on the job, and closely monitoring student attainment can facilitate this communication. Educators also have a responsibility to parents, employers, and the larger community to demonstrate their effectiveness at preparing students who are ready to work successfully.

Levels of Aggregation

Measures of performance can be reported at a variety of levels of aggregation. Data for individual students are the lowest level of aggregation. Student data can be added together and reported for individual classes, courses, programs, schools and institutions, districts, regions, or the state as whole. For states that maintain individual student record systems, aggregation poses no special problems; data may be aggregated in whatever fashion is desired. In effect, the nature of the analysis can determine the appropriate level of aggregation. For states without student record systems, however, the level of aggregation must be determined at the time of developing the accountability system.

Initially, a state's decision about aggregation will probably be determined by the design of its current data system. If, for example, states maintain data on student achievement at the school level (that is, an average score for all students in the school), then the measures of learning gains will initially

have to be developed around this particular data item. Over time, however, states should be able to make changes in their data systems to improve accountability. What level of aggregation, therefore, should be sought?

Student-level data, of course, are ideal. However, if individual student records are not possible, then data for individual programs within an institution are probably the appropriate goal. Suggestions on how to define programs are included in Chapter 2.

For Whom Should Data be Collected?

For purposes of accountability, whose performance should be monitored? That question is not so easily answered, and the law provides no specific instructions. At one extreme, some argue that only those students served by federal vocational education funds need to be included in the system. At the other extreme, the systems could include all students participating in vocational education. Between these two extremes are a variety of other possibilities including

- completers of vocational programs,
- students enrolled in occupationally specific courses,
- students completing more than a specified number of vocational courses,
- students declaring a vocational or career objective.

As a general rule, we favor coverage that is as broad as possible.

Chapter 2 explores these possibilities in more detail. As a general rule, however, we favor coverage that is as broad as possible. Program improvement is a worthy goal for vocational education regardless of the source of funding. Moreover, vocational education seeks to appeal to a wide range of students and bears a responsibility for demonstrating that it is serving them well.

Whatever a state ultimately decides about whom to include in the accountability system, it is important to adopt clear, objective definitions. Preferably, data on course-taking should be used to identify those students whose performance will be monitored. Wherever possible, reliance on student self-reports should be avoided, as these reports are typically inconsistent and unreliable.

Ideally, accountability for vocational education would be well-integrated with a larger system for accountability generally in secondary and postsecondary education. The performance of all students would be monitored and the effectiveness of the entire curriculum regularly evaluated. All students would have career objectives, and it would not be necessary to distinguish vocational students from other students. A few states are pursuing such a model, and their progress will bear close attention.

Setting Standards

Most states will be charting new territory with the development of outcome-based performance standards.

While many states have some experience with monitoring program performance, relatively few have developed statewide performance standards that address learning gains or labor market outcomes. On the contrary, where standards have existed at all, they have been limited to standards about procedures or programmatic inputs (e.g., class size, teacher certification, course completion requirements, and institutional accreditation). As useful as these standards may be, they are not outcome-based, and most states will be charting new territory with the development of performance standards. How might they go about it?

There are at least four ways to start. First, states can set standards relative to existing state averages. Second, they can attempt to identify best practices in exemplary programs or institutions and seek to raise others to the standards of these exemplars. Third, states can seek advice from industry. Fourth, they can examine existing state certification standards where they have already been developed for selected industries or occupations such as cosmetology or nursing. These four are not mutually exclusive, and it is quite possible to pursue all four simultaneously.

Departure from State Averages

Perhaps the easiest approach to setting standards relies simply on the data generated by the accountability system for each of the performance measures. Standards are set relative to statewide averages on these measures. For example, suppose that the state average for placing students in further education, the military, or the labor force is seventy-five percent. This average could become the standard for all programs. Similarly, the standard could be set at some multiple of the state average—1.1, for example, resulting in a standard of 82.5 percent in the case of this hypothetical placement rate.

While this approach has the advantage of simplicity, it suffers from several disadvantages. First, average performance is not necessarily desirable performance. Second, even if standards are set at a multiple of the average, the multiple may or may not be appropriate. Third, in the absence of clear rationales, this approach is quite arbitrary and subject to a variety of political and legal challenges.

In short, this approach can be a useful way to begin the process of setting standards. However, it needs to be supplemented with more research and grounded in educational and industry practices.

Best Practices

Another strategy for setting standards is to examine performance at institutions that operate programs that are widely regarded as exemplary.

How well do the best of best perform? Pursuing this strategy depends on two considerations. First, what constitutes an exemplary program or institution? Second, how widely should one cast the net—within the state, within regions of the country, or nationally?

There are no precise answers to either question. Identifying exemplars may rely on rigorous research, assessments of others knowledgeable about a wide range of programs and institutions, or simple guesswork. To the extent that clear measures of performance have been identified, whether a program is truly exemplary can be put to a simple test: does it significantly outperform other programs on the measures in question? How widely one looks for exemplars is largely a question of resources, time, and data availability. The critical issue in using this best practices approach is understanding why a program achieves outstanding results. Is it clearly a result of *programmatic* features that can be duplicated, or is it more likely the consequence of unique circumstances or the special characteristics of the participants? If programmatic features are the likely explanation, then the performance of the program is truly something to emulate and hold up as a standard to which others should aspire.

Industry Standards

Industry is potentially the best source of information on standards. Standards, after all, are supposed to reflect the levels of performance that are required to perform successfully on the job. Stated differently, the measures and standards adopted by states must ultimately have validity in the workplace. Consequently, industry should be an ideal testing ground for the validity and appropriateness of measures and standards.

Industry is not always able to articulate clearly what is required on the job, and employers do not always have the long-term interests of employees in mind.

Industry, however, is not always able to articulate clearly what is required on the job. Vocational educators have become quite accustomed to the refrain from some employers claiming "Give us people who can read, write, and compute, and we will do the rest." In fact, *other things being equal*, employers will almost always choose candidates with some occupationally specific training over those with none. In addition, industry can be rather shortsighted about the knowledge and skills required. Employers do not always have the longer range interests of employees in mind. In the short run, they may be inclined to settle for a rote understanding of the skills needed immediately in place of more generalized skills that will enable employees to adapt to changes in technology and other working conditions that they are sure to encounter over time.

It is important, therefore, that industry be viewed as an equal partner in the development of measures and standards, not as a higher level authority. Because they are in a better position to take a longer view, educators have much to offer employers, and they should not assume that industry should dictate what is taught or what is expected of students. With these caveats, industry can play an important and lasting role in setting, evaluating, and revising standards of performance.

State Standards

Finally, the development of measures and standards should consider state policies that are already in existence. Almost all states have developed certification standards for some occupations, and in a few instances national standards exist. Occupations that are often state licensed include barbers and beauticians, a wide variety of health occupations, real estate, contracting and building trades, and certain mechanics and repairers. Additionally, there are health and safety standards for many industries. At a minimum, these standards should be reviewed and used as a departure point for discussions about standards. Existing standards may be judged too low or too narrow, but at a minimum they will serve as concrete examples for discussion.

Adjustments and Incentives for Special Populations

Section 115 of the Perkins Act specifies that the accountability systems will include "incentives and adjustments that are designed to encourage service to targeted groups or special populations." The law, however, does not indicate what these incentives and adjustments should be. Chapter 5 provides some specific examples of how this requirement might be met. Here we offer some general guidelines.

"Each system . . . shall include . . . incentives and adjustments that are . . . designed to encourage service to targeted groups or special populations."

PERKINS ACT
Section 115

As a general rule, the same measures and standards should apply to all students. Everyone must compete in the same labor market, whether viewed locally, nationally, or globally. To accommodate the needs of members of special populations, adjustments should first be made to curriculum, pedagogy, and assessment instruments to help them achieve general standards of performance. Only when it is certain that such adjustments are ineffective for some students with special needs, for example, the severely mentally retarded, should standards be lowered or measures changed.

Having stated this general rule, we also recognize that the workplace harbors traditions and mindsets that unnecessarily limit employment opportunities for some special needs students. Special education professionals have been particularly successful at demonstrating that conventional job descriptions can be restructured to create opportunities for students with special needs to work productively.

In short, the primary emphasis should be on high expectations for all students. However, there will always be extreme circumstances that warrant modification and special efforts to accommodate some students. Nevertheless, extreme caution should be exercised in lowering standards, and clear, valid justifications for doing so should be required.

Secondary versus Postsecondary Considerations

Because vocational education has its roots in the high school, federal and state policies tend to overlook the enormous growth of vocational education at the postsecondary level during the past twenty years. Consequently, policy often unwittingly reflects a secondary view of vocational education that is not appropriate for postsecondary offerings. In some instances, this perspective simply fails to understand differences in mission between secondary and postsecondary education; in others, it does not recognize differences in operations or the characteristics of students served.

The design of useful accountability systems for postsecondary vocational education programs poses some special challenges. Postsecondary vocational education serves a much more diversified group of students than do high schools. Attendance at postsecondary institutions is voluntary and of much more varied duration. Students' objectives range from satisfying very immediate, short-range job objectives to long-range degree and career interests. Many students are also likely to attend intermittently over long periods of time.

Therefore, accountability can assume different meanings at the postsecondary level. For example, because attendance is voluntary, it can be argued that postsecondary institutions are more market driven than the secondary schools. Consequently, to survive, they must be more attentive to students' interests and needs. Similarly, they are more likely to maintain close contact with employers and their requirements. Additionally, because postsecondary education ideally builds on a strong elementary and secondary foundation of knowledge and skills, continued learning can be more easily taken for granted. To the extent such assumptions are true, the need for additional accountability provisions is diminished.

While the differences between secondary and postsecondary education are real, one should not exaggerate the differences.

While the differences between secondary and postsecondary education are real, one should not overlook the similarities, especially for the younger population of eighteen- to twenty-two-year-old students who have limited work experience and are not yet sophisticated consumers or investors. For this group, at least, society has a clear interest in accountability that is not much different from what it expects of high schools. One should not exaggerate the differences.

As a practical matter, however, there are circumstances that will make compliance with the accountability provisions of Perkins especially problematic for postsecondary institutions. Measuring learning gains, for example, will pose a real problem in many states. Unlike secondary education, postsecondary education does not periodically assess academic achievement with standardized, statewide tests. Occupational competency testing is even less prevalent.

Additionally, insofar as maintaining and reporting performance outcomes for members of special populations is a feature of state accountability, this requirement may prove troublesome for postsecondary institutions in many

states. While many community colleges provide supplementary services for members of special populations, they are much more likely to rely on student self-assessment of special need. Consequently, procedures for identifying special populations are less uniform and routine.

The uniqueness of postsecondary education should not be overemphasized. Postsecondary accountability is no less important than that for secondary. Nevertheless, there are real differences, and consequently, provisions for phasing-in various aspects of accountability may be especially important at the postsecondary level.

Conclusion

As this first chapter has suggested, there are a number of important considerations affecting the development of accountability systems for secondary and postsecondary vocational education. Some of the issues can become quite technical and complex. Yet it is important not to become bogged down in the complexity. First and foremost, measures and standards should simply be *indicators* of program performance. As indicators, they need not be extremely precise, but they should act as flags for more systematic investigation into what is working, what is not working, why, and why not. They should serve to highlight possible problems that on closer inspection may have perfectly reasonable explanations or may require corrective action. In short, they are ballpark estimates about how programs are doing—the kind of information that any performance-oriented educator would want to know.

It is, of course, true that traditionally educators have not been very performance-oriented. An important corollary, therefore, of developing accountability systems is training teachers and administrators how to use them. Accountability systems are first and foremost a tool for local program improvement. The systems must be keyed to the needs of local level users, not to the needs of high level elected or administrative decision-makers.

Finally, measures and standards are but one tool in the wide array of instruments that must be brought to bear on program improvement. They are no substitute for the broader processes of consultation, monitoring, and management. Relying solely on a rigid system of accountability is certain to lead to gaming, local embarrassment, and distractions from good local assessment and problem solving. States' investment in accountability must be ever mindful of this broader context. Precisely addressing all of the issues posed by developing accountability systems is sure to be an exercise in frustration. Rather than seeking precision and perfection, system developers will best serve the interests of vocational education and the education system at large by keeping systems simple and understandable. Simplicity and clarity, we realize, may pose the biggest challenge of all.

Measures and standards are but one tool in the wide array of instruments that must be brought to bear on program improvement.

2. Getting Started

This chapter discusses a number of key issues that should be considered when designing a system of performance measures and standards, including phasing in a viable system, state and local responsibilities, defining the state's role, scope of the system, defining vocational programs and vocational students, determining the appropriate frequency for performance measures, defining a student cohort, annual evaluations and accountability, responsibility for aggregating and analyzing the data, and establishing a student record system.

Compliance, Viability, and Phasing in the System

As stated in Chapter 1, a careful reading of the Carl D. Perkins Vocational and Applied Technology Education Act of 1990 and a thoughtful interpretation of its provisions can quickly lead to a fairly complex system of performance measures and standards. Consequently, this guide attempts to describe the minimum effort required by law as well as what is needed to create a truly viable system that provides useful and meaningful information on vocational education. However, given that a system of performance measures and standards must be in place for the academic year 1992–1993, it is clear that an ambitious system will have to be phased in over time. Even a minimally compliant system may require several years until all aspects are in place and provide useful information. This guide frequently returns to the idea of phasing in different components of the system of performance measures and standards and provides specific examples for doing so. Furthermore, Chapter 6 on implementation issues discusses in detail the phasing-in process.

State and Local Responsibilities

The Perkins Act places ultimate responsibility for developing the system of performance measures and standards with the State Board of Vocational Education.[1] The State Board is responsible for specifying the core measures and standards to which local schools and institutions will be held accountable, for defining program quality criteria, for conducting an assessment of vocational program quality using these criteria, and for providing technical assistance and program activities to local schools and institutions that do not make sufficient progress in meeting performance

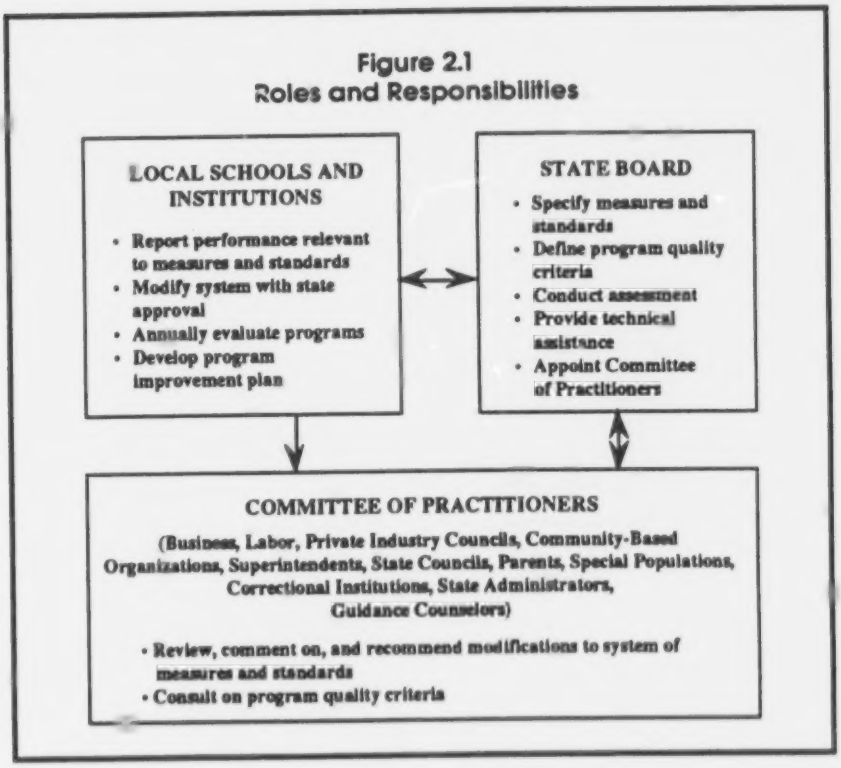
> ***"Each State Board . . . shall develop and implement a statewide system of core standards and measures of performance . . ."***
>
> ***PERKINS ACT***
> ***Section 115***

[1] This guide is intended primarily to provide assistance to those individuals who participate at the state level in the process of developing a system of performance measures and standards. These individuals include members of the State Board of Vocational Education, the State Committee of Practitioners, and those who are staff to these groups.

measures and standards one year after implementing a local program improvement plan. In addition, each State Board must appoint a Committee of Practitioners that will review, comment on, and recommend modifications to the system of performance measures and standards developed by the board and that will provide consultation on the program quality criteria.

Although ultimate responsibility for developing the system of performance measures and standards lies with the state, local schools and institutions participate in developing the system through the Committee of Practitioners. Furthermore, local schools and institutions are responsible for tracking and reporting on the performance of vocational students relevant to the measures and standards specified by the state, for modifying (with approval of the state) the measures and standards based on local conditions,[2] for conducting an annual evaluation of vocational programs based on the measures and standards, and for developing a local improvement plan if they do not make substantial progress in meeting the measures and standards (Figure 2.1).

Figure 2.1
Roles and Responsibilities


LOCAL SCHOOLS AND INSTITUTIONS
• Report performance relevant to measures and standards
• Modify system with state approval
• Annually evaluate programs
• Develop program improvement plan
STATE BOARD
• Specify measures and standards
• Define program quality criteria
• Conduct assessment
• Provide technical assistance
• Appoint Committee of Practitioners
COMMITTEE OF PRACTITIONERS
(Business, Labor, Private Industry Councils, Community-Based Organizations, Superintendents, State Councils, Parents, Special Populations, Correctional Institutions, State Administrators, Guidance Counselors)
• Review, comment on, and recommend modifications to system of measures and standards
• Consult on program quality criteria

[2] Section 115 of the Perkins Act Amendments states that "eligible recipients may make local modifications to such system based on economic, geographic, or demographic factors, or the characteristics of the population to be served. Such modifications shall conform to the assessment criteria contained in the State plan."

The role of the Committee of Practitioners is crucial in developing any state system of performance measures and standards. Through the Committee, the State Board can work with representatives from local schools and institutions who will be held accountable for vocational student performance, individuals from business and industry who will employ vocational students, other professionals and community members who serve vocational students or specific subpopulations of them, and ultimately vocational students themselves. The active participation of all members of the Committee of Practitioners will ensure that the measures and standards that are ultimately chosen are feasible and reflect the concerns, priorities, and needs of all affected parties. Therefore, although the Perkins Act simply requires that the State Board of Vocational Education provide information on the system of performance measures and standards to the Committee of Practitioners and respond to their recommendations, in order to develop a truly worthwhile system, the State Board should work closely with the Committee to ensure that all parties with an interest in vocational education have a voice.

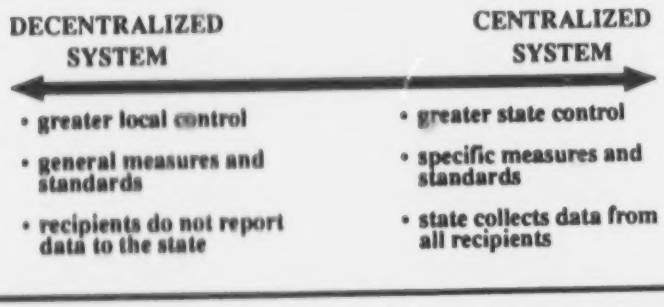
Defining the State's Role

Although the Perkins Act states clearly that the State Board is responsible for developing and implementing the system of core performance measures and standards and that the local recipient is responsible for annually evaluating its vocational programs, these legislative requirements leave a great deal of room for defining what the state's role should be. Depending on the circumstances, the role played by the state may fall anywhere along the following continuum.

"The annual evaluation must be based on the standards and measures developed by the State Board . . ."

PERKINS REGULATIONS §403.191(2)

Figure 2.2
Continuum of Centralization


DECENTRALIZED SYSTEM
CENTRALIZED SYSTEM
• greater local control
• general measures and standards
• recipients do not report data to the state
• greater state control
• specific measures and standards
• state collects data from all recipients

In a highly decentralized system, there is a great deal of local autonomy. Although the State Board defines the core performance measures and standards as required by law, it does so in general terms. For example, the State Board may define a measure of academic gains in the following way: "the percentage of vocational students demonstrating gains on a test of basic academic skills." In this type of system, the state does not specify which test

the local recipients should use nor which basic academic skills (i.e., reading, language, math, science, or social studies) should be tested. Instead, each local school district or postsecondary institution determines which testing instrument to use and which skills to assess. Furthermore, in a highly decentralized system, local recipients may assume full responsibility for collecting data on the core performance measures and determining which of its programs require a local program improvement plan. In this case, the state will be responsible only for monitoring that the local recipient is carrying out the evaluation process.

In contrast, a highly centralized system exhibits a much greater amount of state control. The State Board defines the core performance measures and standards in specific terms. For example, the State Board may define a measure of academic gains in the following way: "the percentage of vocational students demonstrating gains on the _____ test of math and reading skills" at the secondary level or "on the _____ test of adult basic education" at the postsecondary level. In this type of system, the state specifies which test the local recipients should use and which basic academic skills such as reading, language, math, science, or social studies should be tested. Each local school district or postsecondary institution uses the same testing instrument and tests the same skills. Furthermore, in a highly centralized system, local recipients report the relevant data on the core performance measures to the state, and the state participates in determining which programs require a local program improvement plan.

Most states will probably find that their systems of performance measures and standards fall somewhere in between the two extremes.

Where a state's system of performance measures and standards falls on the continuum of centralization depends on a number of factors, including state laws, traditions of local control, existing testing or assessment practices, state capacity for collecting and maintaining large amounts of data, and budget constraints. Moreover, most states will probably find that their systems of performance measures and standards fall somewhere in between the two extremes. For example, a state may define its performance measures in very specific terms but leave responsibility for collecting data and determining which programs require a local program improvement plan to the local recipient. On the other hand, a state may define its performance measures in very specific terms at the secondary education level but define them in general terms at the postsecondary education level.

While a decentralized system makes local autonomy its first priority, a high degree of centralization is not inconsistent with an emphasis on local program improvement. In fact, data that are collected at the state level and that are based on consistent definitions and reporting practices can facilitate local program improvement by providing local school districts and postsecondary institutions with information on the performance of vocational programs across the state. Centrally collected data can provide a basis of comparison for determining which vocational programs are more successful or less. When designing their systems of performance measures and standards, states should seek to strike a balance between respecting local autonomy and providing adequate information for evaluation purposes.

Scope of the Performance Measures and Standards System

The Notice of Proposed Rulemaking specifies in part §403.191 that "each recipient [of Perkins funds] shall evaluate annually the effectiveness of the recipient's entire vocational education program, regardless of which particular projects are assisted with those Federal funds."[3] If this clause is retained in the final regulations, then recipients such as school districts and postsecondary institutions will be required to evaluate their entire vocational education programs, including both Title II and Title III programs, regardless of which specific programs receive Perkins funds. For example, if a local educational agency (LEA) receives Perkins funds, but channels those funds into just two out of seven of its schools, then all seven schools will be required to undertake annual evaluations. In addition, if an LEA or postsecondary institution receives Title II funds only, that recipient will be required to evaluate its Title III programs in addition to its Title II programs. Moreover, the proposed regulations specify in part §403.201 that the system of performance measure and standards should apply to the secondary, postsecondary, *and adult education* programs offered by recipients of Perkins funds.

It should be cautioned that the the final federal regulations may alter the scope of the performance measures and standards system. In particular, the final regulations may require that recipients of Perkins funds evaluate only those vocational programs that directly receive Perkins dollars. In addition, the final regulations may not require that the systems of performance measures and standards apply to adult education programs. However, as stated in Chapter 1, the Act does not prevent states from broadening the scope of their performance measures and standards systems beyond federal requirements. Even if the final regulations limit annual evaluations to those vocational programs that directly receive Perkins funds, individual states may choose to extend their systems to include the entire vocational program of recipients of Perkins funds. Furthermore, the most ambitious states may choose to extend their systems to include all school districts and postsecondary institutions, not just those that receive Perkins funds.

The Perkins Act does not prevent states from broadening the scope of their systems beyond federal requirements.

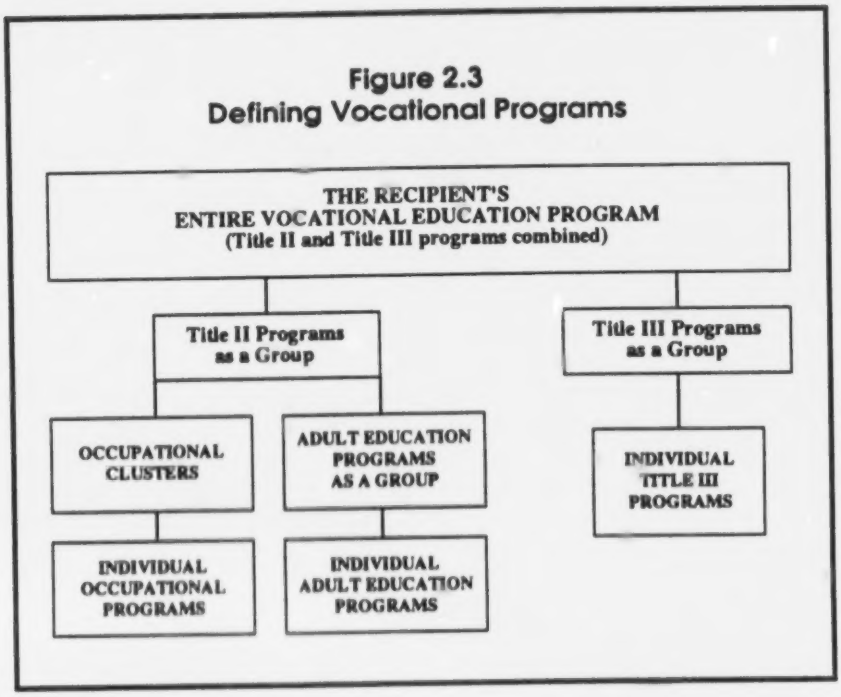
Defining Vocational Programs for Evaluation Purposes

Whether recipients are ultimately required to annually evaluate their entire vocational education program or just those programs that receive Perkins funds, vocational programs must be defined for evaluation purposes. Although an annual evaluation is clearly required, the Act does not specify what form it should take. At what level of detail should states and local recipients evaluate their vocational programs? Should the local recipient evaluate all of its vocational programs as a single group or should it evaluate Title II and Title III programs separately? Should the local recipient evaluate Title II programs in terms of individual occupational programs or

3 "Eligible recipient" is defined in the glossary at the back of this guide.

as occupational clusters? Should adult education programs be evaluated separately from secondary or postsecondary programs? Should Title III programs be evaluated as a group or by individual program? In a centralized system, the state will need to decide up front how to define vocational programs for evaluation purposes and then communicate that definition to the local recipients. In a decentralized system, the local school district or postsecondary institution will develop its own definition. Figure 2.3 illustrates the range of possible definitions that states and local recipients can use.

Figure 2.3
Defining Vocational Programs


THE RECIPIENT'S
ENTIRE VOCATIONAL EDUCATION PROGRAM
(Title II and Title III programs combined)
Title II Programs as a Group
Title III Programs as a Group
OCCUPATIONAL CLUSTERS
ADULT EDUCATION PROGRAMS AS A GROUP
INDIVIDUAL TITLE III PROGRAMS
INDIVIDUAL OCCUPATIONAL PROGRAMS
INDIVIDUAL ADULT EDUCATION PROGRAMS

Each definition of vocational programs has a somewhat different focus. First, choosing to evaluate the entire vocational education program of a school or institution or to evaluate Title II or Title III programs as a group places an emphasis on interschool and interinstitutional comparisons. In this case, the overall performance of a school's Title II programs, for example, would be compared with the overall performance of Title II programs for each school in the state. In contrast, choosing to evaluate vocational programs at a finer level of detail, such as occupational clusters or individual occupational programs, adult education programs as a group or individual adult education programs, or individual Title III programs, enables both inter-school and institutional comparisons and intra-school and institutional comparisons. At this level of detail, vocational program evaluation generates information to help identify stronger and weaker programs within a district, school, or institution, or in the state.

Evaluation of vocational programs at a relatively fine level of detail can provide a powerful source of information for states and local recipients. In addition, this level of detail recognizes that vocational education serves students in many different ways. By evaluating Title III programs separately from Title II programs, or adult education programs separately from secondary or postsecondary programs, states and local recipients can examine the differing impacts on students of different types of vocational education experiences. Furthermore, when defining performance measures and setting relevant standards, the expectations of the impact that vocational education can have on students may differ greatly depending on the type of vocational program in which the student participates.

Evaluation of vocational programs at a relatively fine level of detail recognizes that vocational education serves students in many different ways.

When deciding whether to evaluate Title II programs in terms of occupational clusters or individual occupational programs, states and local recipients should consider the following points. The choice of the cluster definition can facilitate interschool and interinstitutional comparisons, since all the schools or institutions in a state may not offer the same individual occupational programs. In addition, defining vocational programs as occupational clusters also may alleviate problems caused by low enrollment in individual programs when comparing programs within or across schools or institutions. The focus on occupational clusters emphasizes the shift toward educating students in all aspects of the industry that they are planning to enter. However, a strict emphasis on occupational clusters may obscure problems or successes associated with individual programs.

It is possible to phase in different definitions of vocational programs over time. For example, if a state ideally wishes to use a definition of Title II programs based on occupational clusters, but clusters have not yet been developed, then the state can begin in 1992–1993 by evaluating individual occupational programs while setting up a state-level task force to define occupational clusters. Therefore, by 1993–1994, the state may be ready to evaluate programs by occupational cluster. Similarly, states can phase in an evaluation of adult education or Title III programs, beginning with grouped definitions of these programs and moving toward an evaluation of individual programs.

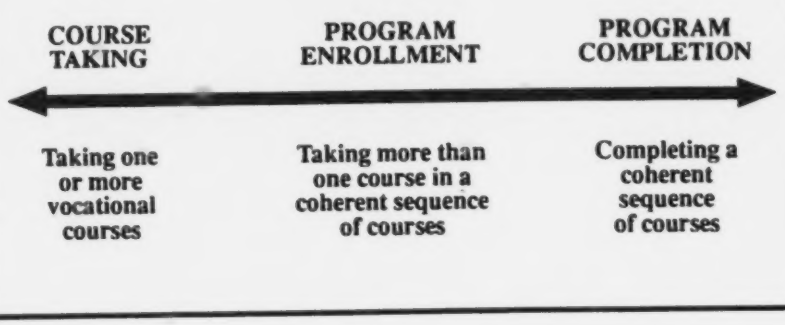
Defining Vocational Students for Data Collection Purposes

After deciding how vocational programs will be defined for evaluation purposes, states (in a centralized system) and local recipients (in a decentralized system) need to define the vocational student population that will be covered by the system of performance measures and standards. Although the Perkins Act requires that annual evaluations be made of vocational education programs, it is the performance of the students in those programs that will actually be measured. Precisely for which students should data be collected? Is a vocational student any student who passes through a vocational program even briefly, or must a student meet more stringent criteria? Should different definitions of vocational students be used depending on the type of vocational program that is being evaluated? Should different definitions be used at the secondary and postsecondary levels?

It is the performance of students in vocational programs that will actually be measured.

Vocational students can be defined broadly as those students who participate in vocational education. However, participation can be further classified into course taking, program enrollment, and program completion. As a result, participation in vocational education can be conceived of as a continuum, as shown in the figure below.

Figure 2.4
Participation in Vocational Education


COURSE TAKING
PROGRAM ENROLLMENT
PROGRAM COMPLETION
Taking one or more vocational courses
Taking more than one course in a coherent sequence of courses
Completing a coherent sequence of courses

The main distinction between course taking and program enrollment or completion is based on whether or not the vocational courses a student takes are part of a coherent sequence of courses. The Notice of Proposed Rulemaking part §400.4(b) defines a coherent sequence of courses as "a series of courses in which vocational and academic education are integrated, and which directly relates to, and leads to, both academic and occupational competencies." If a state chooses to define its vocational student population by either enrollment or completion, it will have to specify what constitutes a coherent sequence of courses for each vocational program area. With respect to Title II occupational programs, some states may have already begun the process of defining coherent sequences of courses, particularly at the postsecondary level. If specified sequences of courses do not currently exist for occupational programs, states can phase in such sequences beginning with the vocational courses that are a recognized part of a vocational program and then add in the academic requirements at a later date. However, the concept of a coherent sequence of courses could be particularly problematic for some Title III programs and for adult education programs. When evaluating these types of vocational programs, states and local recipients may want to stick to a definition of vocational students based on course taking.

Many students at both the secondary and postsecondary levels take vocational courses without necessarily enrolling in or completing a coherent sequence of courses (Hoachlander et al., 1992).[4] Moreover, different levels of course taking, program enrollment, and even program completion are possible. With respect to course taking, one state may collect data on all students who take at least one vocational course, while another may collect data on only those students who take a moderate number of courses, and still another may restrict data collection to only those students who are considered to be "vocational concentrators." With respect to program enrollment, one state may collect data on students who are just beginning a coherent sequence of courses, whereas another may collect data on students who continue well into the sequence. Finally, various states may find different definitions of program completion to be the most appropriate. Tables 2.1 and 2.2 on the following pages illustrate the range of possible definitions of vocational students for both secondary and postsecondary education. However, these are given only as examples; many other definitions would be appropriate.

[4] Fully 97.8 percent of 1987 public high school graduates completed at least one vocational course in high school, and 78.7 percent of all 1980 high school seniors who enrolled in public two-year institutions by 1984 attempted at least one vocational course at the postsecondary level.

Table 2.1
Defining Vocational Students at the Secondary Level

By Course Taking

- Students who take at least one course in any type of vocational education, including Title II and Title III programs.
- Students who take at least one vocational course in a specific occupational area but not necessarily as a part of a coherent sequence of courses.
- Students who take at or above a certain number of courses or Carnegie units (for instance eight units) in any type of vocational education, including Title II and Title III programs.
- Students who take at or above a certain number of courses or Carnegie units (for instance four units) in a specific occupational area but not necessarily as a part of a coherent sequence of courses.

By Program Enrollment

- Students who take at least two vocational courses in a coherent sequence of courses in any type of vocational education, including Title II and Title III programs.
- Students who take at least two vocational courses in a coherent sequence of courses in a specific occupational area.
- Students who take at least two vocational courses in a coherent sequence of courses in a specific occupational area, plus at least one of the advanced academic courses that are a part of the coherent sequence of courses.

By Program Completion

- Students who complete the vocational courses in a coherent sequence of courses in a vocational program area but do not necessarily graduate from high school.
- Students who complete both the vocational and academic courses in a coherent sequence of courses in a vocational program area but do not necessarily graduate from high school.
- Students who complete the vocational courses in a coherent sequence of courses in a vocational program area and graduate from high school.
- Students who complete both the vocational and academic courses in a coherent sequence of courses in a vocational program area and graduate from high school.

Table 2.2
Defining Vocational Students at the Postsecondary Level

By Course Taking

- Students who take at least one vocational course.
- Students who take at or above a certain number of courses or credits, for example twelve credits, in one or more vocational program areas, although not necessarily as part of a coherent sequence of courses.
- Students who take fifty percent or more of all their courses in one or more vocational program areas, although not necessarily as part of a coherent sequence of courses.

By Program Enrollment

- Students who take two or more sequential vocational courses in a single vocational program area.
- Students who take at or above a certain number of vocational courses or credits, for example twelve credits, in a single vocational program area, whether or not those courses are taken sequentially.
- Students who take fifty percent or more of all their courses in a single vocational program area, whether or not those courses are taken sequentially.
- Students who take two or more sequential vocational courses in a single vocational program area plus at least one of the required academic courses that are part of the coherent sequence of courses.

By Program Completion

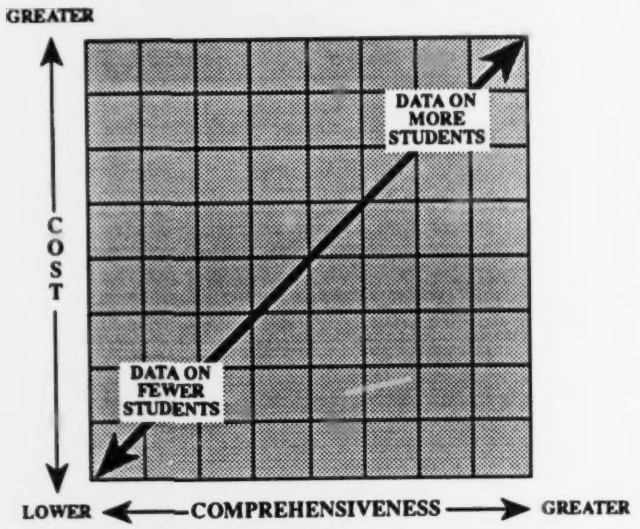
- Students who complete the vocational courses in a coherent sequence of courses, whether or not they complete a certificate or associate's degree.
- Students who complete both the vocational and academic courses in a coherent sequence of courses, whether or not they complete a certificate or associate's degree.
- Students who complete the vocational courses in a coherent sequence of courses and who complete a certificate or associate's degree.
- Students who complete both the vocational and academic courses in a coherent sequence of courses and who complete a certificate or associate's degree.

Other Considerations

- Any of the above for full-time students only.
- Any of the above for both full- and part-time students.

How does a state decide which definition of vocational students to implement? Clearly, how a state chooses to define its vocational student population will have implications for both the cost of collecting data and the comprehensiveness of that data once they are collected (Figure 2.5).

Figure 2.5
Tradeoffs in Defining Vocational Students



On the one hand, the fewer students there are whose performance has to be tracked and reported, the less costly the data collection process will be. On the other hand, the more students there are for whom performance data are available, the more comprehensive a picture of vocational education activities will be possible. For instance, if data are collected on the performance of all students who take at least one vocational education course, then, presumably, the data on those students can be grouped and regrouped, so that states and local recipients can compare the performance of students who take fewer vocational courses with those who take more, the performance of students who take courses in a coherent sequence with those who do not, as well as the performance of completers with noncompleters. Since the vocational education enterprise serves students in many different ways, a system of performance measures and standards that allows for flexibility of data will permit a much more comprehensive evaluation of how well vocational programs meet the varying needs of students.

Since vocational program completion can rightly be considered more of an educational outcome than a descriptive variable, most states will probably choose to define their vocational student populations by course taking or by program enrollment patterns rather than by completion. However, the

definition of vocational students may vary by performance measure. For example, states may want to collect data on the academic gains of all students who enroll in a vocational program, while collecting data on labor market placement for only those students who completed a vocational program.

The process of determining which vocational students will be covered by the system of performance measures and standards is a clear candidate for a phasing in approach. States may want to begin with whatever data system may already be in place and then slowly add on layers of information. For instance, secondary schools might begin by collecting data on students who are enrolled in vocational education programs (students who take more than one course in a coherent sequence of courses), and then move toward including all course takers (students who take at least one vocational course). In addition, postsecondary institutions might begin by collecting data on full-time students only and then move toward including part-time students as well.

Avoiding Data Overkill

Although the Perkins Act requires that recipients annually evaluate their vocational education programs based on the state's core performance measures and standards, it is not necessary that local recipients collect data on all measures for all vocational students every year. Instead, states can keep to a manageable amount the data that must be collected every year by determining the appropriate frequency for performance measures and by defining a student cohort that will be evaluated annually.[5]

Determining the appropriate frequency for performance measures

Although Sections 115 and 117 of the Perkins Act require that annual evaluations be based at a minimum on a measure of academic gains and on one other performance measure, the Act does not require that evaluations be based every year on the full set of core performance measures and standards developed by the state. Consequently, as long as annual evaluations include a measure of academic gains and one other of the performance measures specified in Section 115, states and local recipients have some leeway in deciding which performance measures should be used for evaluation in a given year.

At a minimum, annual evaluations must be based on a measure of academic gains and one other performance measure as specified in Section 115.

However, as cautioned in Chapter 1, when limiting the number of performance measures used for evaluation in a given year, states should be careful to weigh the burden of data collection against the creation of a sound system that reflects the goals of vocational education. In fact, in some cases,

[5] In addition, the final federal regulations pertaining to the Perkins Act may allow states to use a sampling technique that restricts annual evaluations to a sample of vocational programs rather than requiring that the recipient's entire vocational education program be evaluated every year.

states or local recipients may want to collect data more frequently than once a year. However, for practical reasons, states may find it desirable to collect data on certain measures less frequently. For example, a state may choose to conduct a survey of the labor market participation of former vocational students once every two or three years instead of annually in order to reduce the cost of data collection. Although the survey would collect data on all vocational completers who graduated in the previous two to three years, by limiting the frequency of data collection on this placement measure, the state could the reduce administrative costs associated with conducting the survey.

Defining a student cohort to be evaluated annually

A cohort is a group of individuals who have a statistical factor in common.

In addition to restricting the frequency with which performance measures are used for evaluation, states can keep data collection to a manageable level by defining a student cohort that will be evaluated annually. A cohort is a group of individuals who have a statistical factor in common. In the context of performance measures and standards, a cohort could be defined at the secondary level by year in school (that is, senior, junior, sophomore, or freshman) or at the postsecondary level by year of entry into the institution (that is, academic year 1992–1993). By annually reporting on the performance of high school seniors or on the performance of students who entered a postsecondary institution three years prior to the current academic year, all students will eventually be included in the reported cohort. Adjustments can be made as needed for students who dropped out of or completed school before their cohort was reported.

When deciding which cohorts to use, states may want to refer to other educational and employment training programs in the state to see what methods they use for reporting performance. Although it may not always be possible or appropriate, an attempt to be consistent with other programs could reduce some of the potential confusion, particularly for adult education schools and postsecondary institutions that might have to report on the performance of their students to several different agencies.

Defining a student cohort poses more problems at the postsecondary level. While the majority of students can expect to spend four years in high school, attendance at postsecondary institutions is much more varied. Many students attend only part time, take remedial courses either before or along with other classes, "stop out," and re-enter for a variety of reasons including taking time out to gain work experience. Some do not complete a certificate or associate's degree because their education objectives do not include such completion. Consequently, states and postsecondary institutions must be creative about defining an appropriate cohort at this level. The most workable alternative seems to be to define the cohort by starting date, and then to choose a critical point that represents the expected completion of the majority of full-time students. It may be necessary to establish a separate expected completion period for part-time students.

How does the reporting of cohorts work? If seniors are chosen as the cohort at the secondary level, then schools will annually report all information relevant to the state's performance measures for that year's seniors. For example, at the end of the 1992–1993 academic year, schools will report on the performance of the class of 1993 (Table 2.3). However, this does not mean that all relevant data had to be collected during 1992–1993. If, for instance, tests of academic skills are administered in the ninth and eleventh grades, then schools will report the eleventh-grade attainment for the class of 1993, as well as the ninth- to eleventh-grade gain for the class of 1993. Similarly, at the end of the 1993–1994 academic year, schools will report the eleventh-grade attainment for the class of 1994, as well as the ninth- to eleventh-grade gain for the class of 1994. By defining a single cohort for reporting purposes, annual reporting describes a group of students (in this case, seniors), rather than describing the data collection activities for a given year.

Table 2.3
Annual Reporting by Cohort

Reporting date:	**Summer 1993**	**Summer 1995**
Level:	**Secondary**	**Postsecondary**
Cohort:	**1992–1993 Seniors**	**1992–1993 Entrants**
	• Eleventh grade academic attainment	• Program completion rate
	• Ninth–Eleventh grade academic gain	• Basic academic skills gains

At the postsecondary level, if the cohort is defined as full-time students in the institution three years prior to the given academic year, then institutions will annually report all information relevant to the state's performance measures on that group of entrants (Table 2.3). For example, after the end of academic year 1994–1995, institutions will report all available information on part-time students enrolled during the academic year 1992–1993. If the expected completion period for part-time students is five years, then after the end of academic year 1996–1997, institutions will report all available information on part-time students enrolled during the academic year 1992–1993, as well as all available information on full-time students enrolled during the academic year 1994–1995.

In some instances, however, states may want to make some exceptions to the cohort rule (Table 2.4). For example, if cohorts are defined by expected completion, this does not leave enough time to measure post-program outcomes, such as placement in employment or further education. To address this problem, secondary schools might annually report on placement of the previous year's seniors, while reporting on the performance of the current year's seniors relative to all other measures. For example, at the end of the 1993–1994 academic year, schools will report on placement measures for the class of 1993, while reporting on all other measures for the class of 1994. Similarly, at the postsecondary level, institutions might annually report on placement of students who enrolled four years prior to the current year, while reporting on the performance of students who enrolled three years prior to the current year relative to all other measures. For example, at the end of the 1995–1996 academic year, institutions will report on placement measures for students entering the institution during the 1992–1993 academic year, while reporting on all other measures for students entering the institution during the 1993–1994 academic year.

Table 2.4
Exceptions to the Cohort Rule

Reporting date:	**Summer 1994**	**Summer 1996**
Level:	**Secondary**	**Postsecondary**
Cohort:	**1992–1993 Seniors**	**1992–1993 Entrants**
	• Labor market placement rate	• Labor market placement rate
Cohort:	**1993–1994 Seniors**	**1993–1994 Entrants**
	• Eleventh grade academic attainment	• Program completion rate
	• Ninth–Eleventh grade academic gain	• Basic academic skills gains

Other exceptions to the cohort rule may also be appropriate. For instance, at the secondary level, states may decide it is important to report on the performance of all special education students every year, or on all students performing below expectations on learning gains, in both cases not just those

who were seniors. Similarly, at the postsecondary level, states may decide it is important to report on the performance of all students who entered the institution with insufficient basic skills, rather than just those below-basic students who were part of that year's reported cohort.

Clearly, by choosing a cohort based on expected completion, data collection and reporting will have to be phased in over time. At the secondary level, some measures, such as program completion, retention in school, or occupational competency attainment, can probably be reported for the class of 1993 at the end of the 1992–1993 academic year. However, if only a single academic skills test is offered to students in the state, for instance in the twelfth grade, then only academic skills attainment and not academic skills gains can be reported for the class of 1993. However, during the 1992–1993 academic school year, a pretest could be administered for the first time to eleventh graders, so that, in conjunction with the usual twelfth-grade test, both academic attainment and gains could be reported for the class of 1994 at the end of the 1993–1994 academic year. Similarly, schools will have to wait until the end of the 1993–1994 academic year to report the placement rates for the class of 1993. However, phasing in procedures will have to be based on a state's specific performance measures as developed in Chapter 3.

Data collection and reporting will have to be phased in over time.

At the postsecondary level, the problem of data availability is particularly acute, because student cohorts are defined by the beginning of a student's time at the institution, rather than by the last year in school, as was done at the secondary level. However, at the end of the 1992–1993 academic year, institutions could certainly report baseline data for their chosen cohort. For example, institutions could report the number of students entering the institution as full-time students during the 1992–1993 academic year who stated an intention to follow a coherent sequence of courses in a vocational program area. In addition, institutions could report other descriptive data on the 1992–1993 entrant cohort as well as the basic academic skills gains made by the end of the first year at the institution for students who tested below an acceptable level upon entry at the beginning of the year. Furthermore, some modified version of program completion could be reported, for example, the percentage of students stating they would follow a coherent sequence of courses in a vocational program area who actually demonstrated taking the first courses in that sequence. Finally, retention in school after one year could be reported for both full- and part-time students. Again, specific phasing in procedures and temporary modifications to chosen performance measures must be based on a state's performance measures, as developed in Chapter 3.

Finally, states may decide not to define a single cohort for annual reporting. Instead, states may prefer to collect the most current data available on individual performance measures. In this case, a distinct cohort would be defined for each performance measure. While the emphasis of a single cohort is on evaluating the effectiveness of vocational education programs for a particular group of students, the emphasis of a multiple cohort approach is on the effectiveness of vocational education programs at a single point in time (Table 2.5).

States may prefer to collect the most current data available for each performance measure.

Table 2.5
The Multiple Cohort Approach

Reporting date: Level:	Summer 1994 Secondary	Summer 1996 Postsecondary
Cohort:	**1993–1994 Juniors**	**1993–1994 Entrants**
	• Eleventh grade academic attainment	• Advanced academic skills attainment
	• Ninth–Eleventh grade academic gain	• Program completion rate
Cohort:	**1993–1994 Seniors**	**1992–1993 Entrants**
	• Program completion rate	• Labor market placement rate
Cohort:	**1992–1993 Seniors**	**1994–1995 Entrants**
	• Labor market placement rate	• Basic academic skills gains

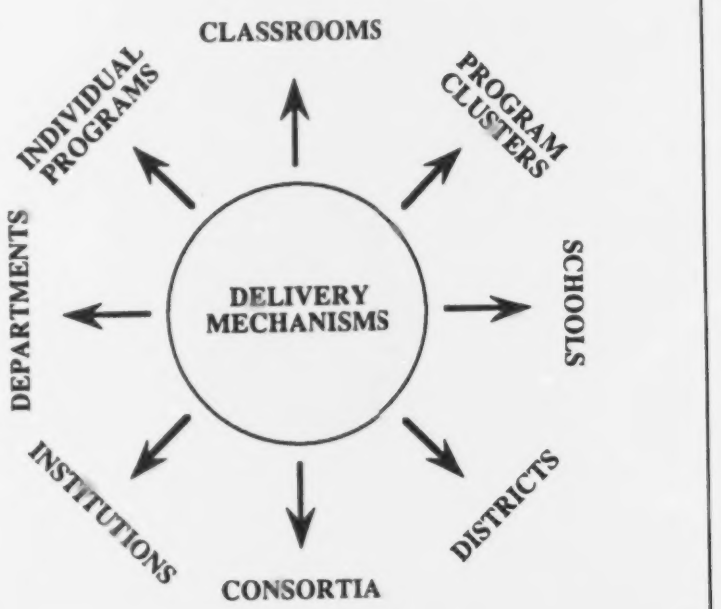
At the secondary level, if tests of academic skills are offered in the ninth and eleventh grades, then a school that takes the multiple cohort approach can annually report the attainment of eleventh graders and the gains made by eleventh graders since the ninth grade. At the same time, schools can report program completion for seniors and placement rates for last year's seniors. At the postsecondary level, institutions can simultaneously report attainment of advanced academic skills and program completion rates for students who enrolled three years prior to the current year, placement rates for students who enrolled four years prior to the current year, and academic skills gains for students completing a posttest of basic academic skills (which, for instance, might be given one year after entrance to all students who failed the pretest of basic academic skills upon entrance). If this method is chosen, states will have to wait until they define their individual performance measures before assigning an appropriate cohort to each measure.

Annual Evaluations and Accountability

As stated previously, the Perkins Act requires that local recipients annually evaluate their vocational education programs and that they develop a program improvement plan if they do not make substantial progress in meeting the state's core standards and measures. However, the Act does not specify the scope of the program improvement plan. Should the local recipient write a single improvement plan that encompasses all of its schools or institutions that offer vocational programs or should it write an improvement plan for each school or institution that does not make substantial progress? Should the local recipient write a single improvement plan that encompasses its entire vocational education program activities or should it write an improvement plan for each program that does not make substantial progress? Should the local recipient write an improvement plan based on some other criteria? Ultimately, states (in a centralized system) and local recipients (in a decentralized system) must decide what level of activity is being held accountable.

Vocational education is offered simultaneously through a number of delivery mechanisms. Consequently, vocational education programs can be evaluated and program improvement plans can be written at any of these levels (Figure 2.6).

Figure 2.6
Vocational Education Delivery Mechanisms



To some extent, the scope of the program improvement plan will be driven by the state or local recipient's definition of vocational programs. If the state defines vocational programs as the recipient's entire vocational program, then the program improvement plan will be written either at the agency level (consortia or school district) or at the building level (school or institution). However, if the state or local recipient defines vocational programs in terms of individual occupational, individual adult education, or individual Title III programs, then the program improvement plan can be written at any level, including the agency level (consortia or school district), the building level (school or institution), or any of the finer levels of detail (program, department, or classroom). Similarly, if the state or local recipient defines vocational programs in terms of occupational clusters, then the program improvement plan can be written at the agency level (consortia or school district), the building level (school or institution), or the cluster level.

The criteria for triggering a program improvement plan can be based on a combination of the above delivery mechanisms. For example, the overall performance of a school or institution is first compared with the overall performance of other schools or institutions in the state, then, if that school or institution performs poorly, it will be required to develop an improvement plan but only for its weaker programs. Mechanisms for triggering the requirement to develop a program improvement plan will be covered in more detail in Chapter 4 on performance standards.

Responsibility for Aggregating and Analyzing the Data

Although local recipients are ultimately accountable for the performance of their vocational programs, it is not necessary that they be responsible for aggregating and analyzing the data on their students and programs. In a centralized system, the local recipient submits data to the state on its individual students, including program affiliation and other necessary descriptive information, and then the state tallies the relevant percentages and averages and reports them back to the local recipient. At the same time, the state informs the local recipient how its performance compared with that of other recipients and whether it is required to develop an improvement plan.

States will have to weigh the advantages and disadvantages of a centralized approach when deciding who should be responsible for aggregating and analyzing the data.

States will have to weigh the advantages and disadvantages of a centralized approach when deciding who should be responsible for different aspects of the reporting process. There are a number of advantages to making the state responsible for aggregating the data including the following: States may have greater automated capacity and more technical personnel than local schools or institutions to handle the necessary data processing and management; states can monitor the quality of the data and ensure that consistent methods were used to calculate performance measures; and states can efficiently provide information to the school or institution on how its performance compared with other schools or institutions. However, there are several disadvantages to making the state responsible for aggregating the data including the following: Some states may not have the automated

capacity or the personnel to process large amounts of data; variability in quality and consistency of data may make the central handling of data an inefficient and time-consuming process; and state handling of data shifts the focus away from local efforts, and places an emphasis on interschool or interinstitutional comparisons, rather than on local evaluation efforts.

Establishing a Student Record System

An important implication of examining the performance of students in vocational programs is that the accountability system should be based on a student record system. At this point, states have theoretically made preliminary decisions about which vocational programs and students will be covered by the system, which measures and cohorts will be reported, and what the scope of the program improvement plan will be. Next, states must address a broad range of issues about what specific data elements should be included in a student record system and what data collection procedures should be followed.

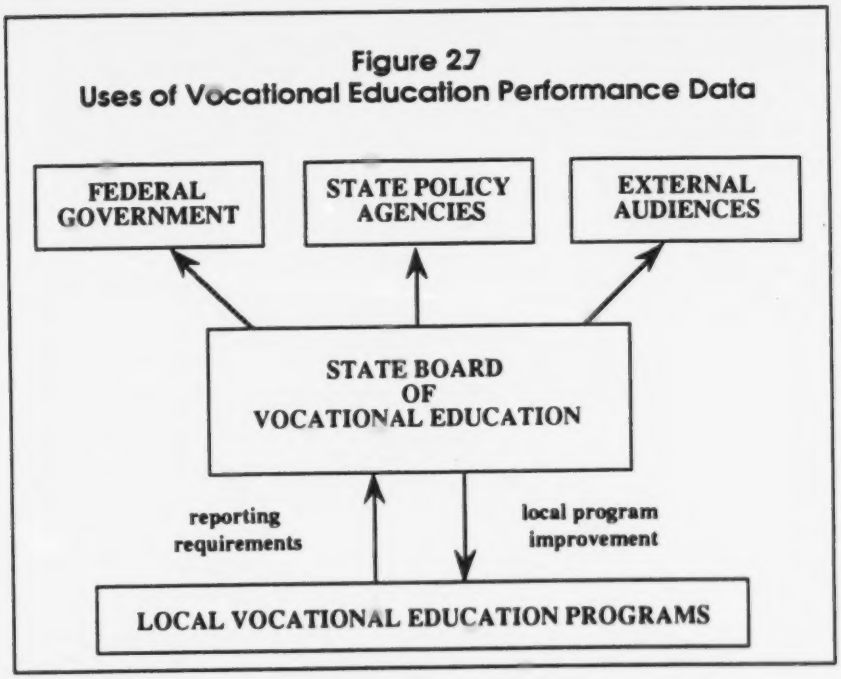
Setting Priorities for Uses of Data

During 1993, statewide data on the performance of students in various vocational programs will begin to be available. How will the data that are collected be used? What uses should be given priority? As stated previously, the Perkins Act requires that local schools and institutions use information gathered on the performance measures and standards to conduct annual evaluations of the effectiveness of their vocational programs. However, it is possible to conceive of additional uses of the data, including meeting federal requirements, providing information to external audiences, and informing state policy. Although all of these uses are legitimate, and indeed some mix of all of them is probably a desirable goal, clear priorities among uses must be set in order to avoid sending conflicting messages that may result in an incoherent system of measures, standards, and reporting practices (Figure 2.7).

An emphasis on local program improvement should remain the guiding principle throughout states' development of their systems of performance measures and standards. Too great an emphasis on meeting federal requirements can contribute to a sense of burden, where local schools and institutions feel that they participate only because they have to. Similarly, too great an emphasis on providing information to external audiences can shift attention away from the place where learning is taking place. Clearly, there needs to be a careful balance between the mandate for accountability beyond the classroom and focused local improvement efforts. Finally, too great an emphasis on uses of data for state policy can create a cumbersome, unfocused system that attempts to do too much.

An emphasis on local program improvement should be the guiding principle throughout the development of a system of performance measures and standards.

Figure 27
Uses of Vocational Education Performance Data



Specifying a Triggering Mechanism for Data Collection

States need to work with local schools and institutions to determine the most appropriate mechanism for triggering data collection on vocational students. For example, since the Perkins Act requires that states measure student gains in basic and more advanced academic skills, a state cannot wait until a student has completed a vocational program to begin assessing that student's academic skill gains. Similarly, if a state decides to measure retention in school, such as high school graduation or postsecondary certificate or degree completion within a specified period of time, schools and institutions need to know where each graduating or completing student was at the beginning of the relevant time period. Clearly, schools and institutions must begin to collect data on vocational students early in the secondary or postsecondary education process. Furthermore, collecting baseline data at the time a student begins to participate in vocational education will facilitate an evaluation of the impact of vocational education on that student.

If a state defines vocational students by course taking, then schools and institutions can begin to collect data on a student as soon as that student takes a single vocational course. If a state defines vocational students by either program enrollment or completion, then schools and institutions can begin to collect data on a student as soon as that student takes the first course in an occupational area or the first two courses in a coherent sequence of courses. Alternatively, data collection on enrollees or completers could be prompted by some combination of taking a first occupational course, or the first two courses in a sequence, and intent to enroll in or complete a vocational program.

Deciding When to Collect Data on Nonvocational Students

The decision to collect some data on all students may also be desirable since vocational students do not fit into a single mold, and early identification of vocational students may not be easy. At the secondary level, some students may take their first vocational course in the ninth grade, while others may wait until the eleventh or twelfth grade. If an academic skills test is offered in the tenth and twelfth grades, those students who took their first vocational course in the eleventh grade might be left out of the first stage of the data collection process, unless the skills test was offered to all students in the tenth grade. However, if the test was offered to all students, some mechanism must exist to identify which students taking the test were vocational students. At the postsecondary level, particularly for part-time students, it might take several semesters until any course-taking pattern is clear. Some sort of ongoing monitoring system should probably be put in place to catch those students who begin to resemble vocational students.

Collecting some data on all students may be desirable since early identification of vocational students may not be easy.

States may also find it useful to compare the performance of vocational students with that of nonvocational students in order to determine the impact of vocational education on student performance.[6] However, it is important to keep in mind that the first priority of the system should be to provide useful information for local program evaluation and improvement. While a highly ambitious system might seek to answer a broad range of interesting questions on vocational student performance, the system should focus first on putting the mechanisms in place for answering basic questions about program effectiveness.

Specifying the Data Elements that Should be Part of the Student Record

As a final step, states should be very specific about what descriptive data elements should be part of the student record. While keeping local program improvement a priority, states should plan ahead for the kinds of analysis they would like to make based on the data collected through the performance measures and standards system. Even at the end of the 1992–1993 academic year, states may begin to evaluate what types of programs appear to be successful for what types of students. In order to ensure that relevant data are available and consistent across a state, states should communicate clearly to local schools and institutions about the kinds of descriptive information they want recipients to collect in conjunction with performance data.

[6] In order to estimate the impact of vocational education on students, it would be necessary to control for different characteristics of students, such as demographics, since the type of student who elects to participate in vocational education might be different from the type of student who elects not to participate. However, it must be cautioned that a true determination of the impact of vocational education programs on students would require random assignment of students to vocational programs, an evaluation method that is beyond the scope of the accountability system required by Congress.

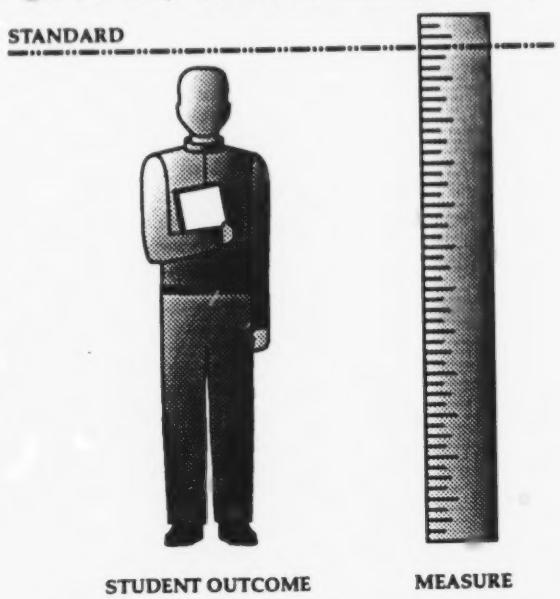
3. Performance Measures

By building on the preliminary system design decisions that were made in Chapter 2, this chapter describes a step-by-step process for defining performance measures and provides examples of learning, labor market, and accessibility measures for both secondary and postsecondary vocational education. The process of defining measures leads to the setting of performance standards that is covered in Chapter 4.

Working Definitions

The Notice of Proposed Rulemaking (NPRM) part §400.4(b) defines measures as "a description of an outcome," and standards as "the level or rate of an outcome." Although the NPRM does not offer a definition, this guide defines an outcome as "a measurable aspect of student performance." Consequently, an outcome indicates the condition that will be measured, while a measure specifies *how* that condition will be measured, and a standard represents the level against which performance on the measure will be evaluated (Figure 3.1).

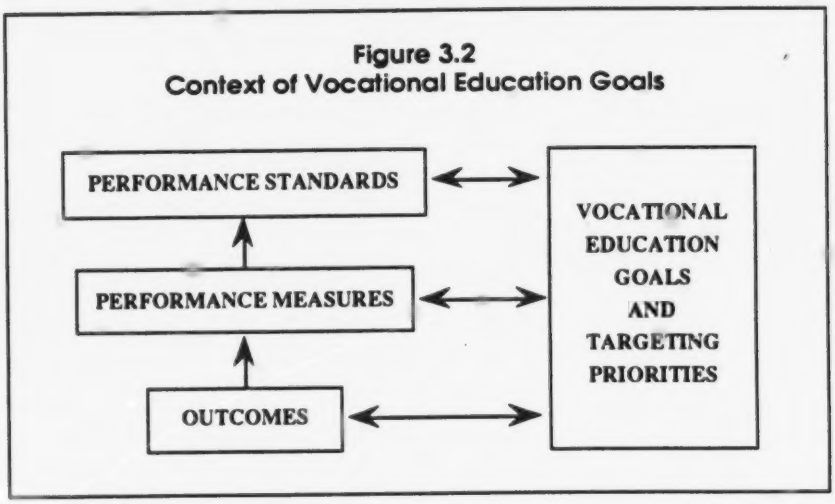
Figure 3.1
Outcomes, Measures, and Standards



Accountability Within the Broader Context of Vocational Education Goals

As a state begins to define its performance measures and standards, it should keep in mind that the accountability system must be developed within the broader context of the state's goals for vocational education. If a state's chosen measures and standards do not fit into its overall goals, the state risks that those measures and standards will replace its goals as the driving force behind vocational education efforts. At a minimum, an ill-conceived system of performance measures and standards can distort educational priorities. The process of developing an accountability system involves three main steps: (1) choosing a mix of student outcomes, (2) defining one or more measures for each outcome, and (3) then setting standards for each measure. Consequently, at each stage in the process, the state must ask whether its chosen mix of outcomes, measures, and standards reinforces the goals and targeting priorities it has set for vocational education (Figure 3.2).

Figure 3.2
Context of Vocational Education Goals



Developing Parallel Systems for Secondary and Postsecondary Education

The Perkins Act and proposed regulations do not directly address the issue of whether or not states should develop a single set of performance outcomes, measures, and standards for both secondary and postsecondary vocational education. However, given the absence of injunctions against developing separate systems and given the practical reasons for doing so, states can assume until otherwise indicated that separate sets of outcomes, measures, and standards are allowable. However, given the emphasis in the Act on coordination of secondary and postsecondary programs, specifically through Tech Prep, states should attempt to promote coordination and consistency while recognizing important distinctions between the two levels of programs.

The Process of Defining Performance Measures

The process of defining performance measures should include the following steps:

1. Choose an appropriate mix of outcomes
2. Decide whether to include process variables
3. Decide how to measure each outcome
4. Specify the chosen measures
5. Translate student measures into program measures
6. Report program measures according to the state's definition of vocational programs

Step 1: Choose an appropriate mix of outcomes

As a first step in defining performance measures, a state should choose a mix of outcomes that both meets federal requirements and reflects state goals for vocational education. The following list of learning, labor market, and accessibility outcomes illustrates the broad range of outcomes from which a state can choose.

This guide uses slightly different language than is used in the Perkins Act for some learning outcomes. The term "competency attainment" used in section 115 of the Act is referred to here as "specific occupational competency attainment." Similarly, the term "job or work skill attainment or enhancement including student progress in achieving occupational skills necessary to obtain employment in the field for which the student has been prepared, including occupational skills in the industry the student is preparing to enter" is referred to here as "general occupational skills attainment or enhancement."

Table 3.1
Student Outcomes

Learning Outcomes

- Gains in basic and more advanced academic skills* or attainment of such skills
- Specific occupational competency attainment**
- Specific occupational competency gains*
- General occupational skills attainment or enhancement** ("all aspects of the industry")
- Employability skills attainment or enhancement ("workplace basics")
- Retention in school or completion of secondary school or its equivalent**
- Vocational course completion
- Vocational program completion

Labor Market Outcomes

- Placement into additional training or education, military service, or employment**
- Time until placement
- Retention in first job
- Earnings (e.g., entry level wage or quarterly earnings)
- Rate of earnings increase
- Employer satisfaction

Accessibility Outcomes

- Participation in vocational education programs, including course taking, enrollment, and completion
- Participation in the labor market, including placement into additional training or education, military service, or employment

* These outcomes are required by law.
** This outcome is one of several from which a state is required to choose one.

A mix of outcomes that draws from all three—learning, labor market, and accessibility categories—will provide a reasonable balance among vocational education objectives.

There are hundreds of possible combinations of the above outcomes. However, states may find it helpful to follow several simple guidelines in choosing the appropriate mix. First, more than the two sets of required outcomes will probably be necessary to develop a well-balanced system. Second, a mix of outcomes that draws from all three learning, labor market, and accessibility categories will probably provide a reasonable balance among vocational education objectives and will prevent an undue focus on one type of outcome to the exclusion of another. Third, the number of outcomes should not become unmanageable, or the system will lack focus and direction. Fourth, a mix of outcomes should include attention to both attainment and gains, in order to provide a balance between serving low-achieving students and setting an acceptable level of performance for all students.

A mix of outcomes that includes both attainment and gains combines the advantages of both approaches while avoiding the disadvantages. A gains approach holds programs accountable for improving the skills of students, thereby providing an incentive to enroll students who demonstrate a relatively low initial level of performance. This approach removes the pressure to cream, that is, to enroll only those students who are most likely to achieve a high level of performance. However, the problem with the gains approach is that it doesn't recognize that all students must achieve a certain level of competency to be successful in the labor market or to further their education. While some students may make significant progress, they may fall short of the entry-level requirements of industry and postsecondary institutions. An attainment approach, on the other hand, expects that students achieve a specified competency level, but risks providing an incentive to cream. By combining the gains and attainment approaches, programs will have the incentive to enroll both students who have a high probability of achieving the specified attainment level and those who will make the greatest gains.

In addition to the four guidelines listed above, states must consider the requirements of Section 117 of the Perkins Act when deciding which outcomes to include in their chosen mix. Although the Act does not specify that states must include the outcome pertaining to general occupational skills (preparation in all aspects of the industry) or any of the accessibility outcomes, states are required to evaluate these aspects of their programs annually.[1] States may find that including outcomes associated with these two areas in their systems of performance measures and standards is an efficient method of collecting information to address Section 117 requirements.

States must consider the requirements of Section 117 when deciding which outcomes to include.

Tables 3.2 and 3.3 illustrate two models of possible outcome mixes. Although each model has an implicit program emphasis and set of targeting priorities, neither represents the single ideal mix of outcomes. Many other variations would also lay the foundation for a sound accountability system.

[1] Section 117 of the Act states that each recipient of Perkins funds shall, as part of the annual evaluation, review programs in order to: (1A) "identify and adopt strategies to overcome any barriers which are resulting in lower rates of access to vocational education programs or success in such programs for individuals who are members of special populations"; (1B) "evaluate the progress of individuals who are members of special populations in vocational education programs assisted under this Act"; and (2) "evaluate the progress of vocational education programs assisted under this Act in providing vocational education students with strong experience in and understanding of all aspects of the industry the students are preparing to enter."

BEST COPY AVAILABLE

Table 3.2
Model #1

Possible Outcome Mix

- Gains in basic and more advanced academic skills*
- Specific occupational competency gains*
- Retention in school**
- Placement into additional training or education, military service, or employment**
- Earnings (e.g., entry level wage)

* These outcomes are required by law.
** This outcome is one of several from which a state is required to choose one.

Model #1 includes three learning outcomes and two labor market outcomes. The model does not explicitly address the issue of accessibility for special populations, although the measures of academic skills and occupational competency are specified in terms of gains that emphasize service to low-achieving students. Although the model includes the required outcomes pertaining to learning gains, the primary focus appears to be on success in making it through school and beyond. The combination of the outcomes of retention in school, placement, and earnings emphasize completion of school and movement into the labor force. Furthermore, the model places little emphasis on direct evaluation of vocational programs. Although placement in related employment and wages earned can be indirect ways of evaluating how well a vocational program has served students, only one direct outcome was used, i.e. occupational competency gains.

Table 3.3
Model #2

Possible Outcome Mix

- Attainment and gains in basic and more advanced academic skills*
- Specific occupational competency gains*
- Specific occupational competency attainment**
- General occupational skills attainment and gains** ("all aspects of the industry")
- Retention in school**
- Vocational program completion
- Placement into additional training or education, military service, or employment**
- Employer satisfaction
- Enrollment and completion of special populations in vocational education programs
- Placement of special populations into additional training or education, military service, or employment

* These outcomes are required by law.
** This outcome is one of several from which a state is required to choose one.

Model #2 includes five learning outcomes, two labor market outcomes, and two accessibility outcomes. In addition to the two accessibility outcomes, the model places further emphasis on serving low-achieving students by including both attainment and gains in its first three learning outcomes. While model #2 includes retention in school and placement after school, as does model #1, it also contains several direct ways to examine the effectiveness of vocational programs, including specific occupational competency, general occupational skills, and vocational program completion. In comparison with model #1, model #2 covers a broader range of concerns, including academic skills, vocational skills, completion (both of school and of vocational programs), and employment (both from the student and employer point of view). Finally, the accessibility outcomes reinforce the model's other concerns, specifically, vocational program completion and placement in the labor market.

While it might appear that model #2 is a better model because it supports a more comprehensive vision of vocational education efforts, model #1 may be the most appropriate under certain circumstances. Specifically, model #1 may serve the needs of a state that has fewer resources and whose vocational education goals emphasize moving out into the labor market. However, it should be cautioned that a state that applies model #1 may encounter difficulty in evaluating the effectiveness of its programs in providing equal access to special populations, as well as in providing all students with experience in all aspects of the industry they are preparing to enter.

A final consideration for states is whether to choose different mixes of outcomes for secondary and postsecondary programs. Different mixes are desirable when secondary and postsecondary education have different goals. For example, in one state, secondary education may focus on providing students with the skills that will enable them to be successful throughout their working lives, while postsecondary education may focus on preparing them for entry into and success in a specific occupation. In this state, the appropriate mix of outcomes for secondary education would include academic, general occupational, and employability skills, while the appropriate mix for postsecondary education would include academic and specific occupational skills and placement in the labor market.

Different mixes of outcomes are desirable when secondary and postsecondary education have different goals.

When specifying measures, states will probably find it necessary to use different measures for secondary and postsecondary education. Even when a state chooses the same mix of outcomes for both the secondary and postsecondary levels, the precise methods used to measure those outcomes may differ. For example, although a state may choose to use retention in school as one of its outcomes for both secondary and postsecondary education, the state will probably measure retention differently at the two levels. However, in some cases, even the precise measures may be the same. For instance, a state could choose to use the same test of adult basic education to assess basic academic skills at both the secondary and postsecondary levels.

Step 2: Decide Whether to Include Process Variables

As was previously mentioned in Chapter 1, the Perkins Act of 1990 places a much stronger emphasis on measuring student *outcomes* than does any previous legislation. Not only do the regulations define measures and standards in terms of outcomes, but also section 115 of the Act requires that the minimum two performance measures be based on a short list of student outcomes. However, section 116 of the Act has a different focus, exhorting states to conduct an assessment of vocational program quality based on measurable *process* criteria that include the following factors:

- Integration of academic and vocational education;
- Sequential courses of study leading to both academic and occupational competencies;
- Increased student work skill attainment and job placement;
- Increased linkages between secondary and postsecondary educational institutions;
- Instruction and experience, to the extent practicable, in all aspects of the industry the students are preparing to enter;
- The ability of the eligible recipients to meet the needs of special populations with respect to vocational education;
- Raising the quality of vocational education programs in schools with high concentrations of poor and low-achieving students;
- The relevance of programs to the workplace and to the occupation for which students are to be trained, and the extent to which such programs reflect a realistic assessment of current and future labor market needs, including needs in areas of emerging technologies;
- The ability of the vocational curriculum, equipment, and instructional materials to meet the demands of the work force;
- Basic and higher order current and future workplace competencies that will reflect the hiring needs of employers; and
- Other factors considered appropriate by the state board.

For purposes of assessing program quality on an ongoing basis, states may choose to mix in process variables with their chosen student outcomes.

In order to collect information on these factors for purposes of assessing program quality on an ongoing basis, some states may choose to mix in process variables ("measurable objective criteria") with their chosen student outcomes. For example, if a state decides that "increased linkages between secondary and postsecondary educational institutions" (i.e., participation in a Tech-Prep program), are an important aspect of quality vocational programs, then the state can add this factor to its list of outcomes. Consequently, the state will need to develop ways to measure this factor just as it will need to develop measures of student outcomes. Alternatively, a state may choose

not to combine process variables with its outcomes and instead will collect information on its program quality criteria through another channel.

States must also decide whether to collect additional process and input information through the system of performance measures and standards for purposes of program evaluation. Specifically, once a state has analyzed the annual data collected on performance measures and standards and has identified which schools or institutions or which programs must develop a program improvement plan, how is that school, institution, or program going to determine what changes need to be made? While information on the above program quality factors may offer some clues to what makes up a successful program, other information, such as student attendance, facilities and equipment, and faculty experience, may provide some insight for program improvement efforts. Once again, states can choose to add this process variable to their list of outcomes, and then develop specific ways to measure it, or they can choose to collect information through a different channel.

Step 3: Decide How to Measure Each Outcome

The next step in the process of defining performance measures is to decide how to measure each outcome chosen under step #1. As stated previously, outcomes indicate the condition that will be measured, while performance measures specify *how* those outcomes will be measured. The first outcome, gains in basic and advanced academic skills, can be measured in different ways such as through the use of a test of basic academic skills, through course completion in an advanced academic area, and through the use of an alternative assessment such as a writing sample or comprehensive project. Similarly, retention in school can be measured in different ways, such as high school graduation, postsecondary degree or certificate completion, or completion of a student's stated education goals.

When deciding how to measure an outcome, a state will be faced with choosing from existing methods of assessing and determining performance and developing entirely new methods. While some performance measures are obviously quantifiable and therefore measurable, particularly those that rely on standardized tests or common measures of performance such as high school graduation or degree completion, many alternative ways of measuring student performance can also be quantifiable. Chapter 7, on assessment, discusses the broad range of assessment instruments available and provides examples of and resources for finding alternative assessments. More qualitative measures, such as writing samples and senior projects or degree projects, are measurable as long as the criteria used in the assessment are explicitly defined. For example, an assessment based on a writing sample should specify the criteria used to evaluate the student's performance. Then, a student could be determined to have passed or failed, or even to have passed with distinction. Similarly, an alternative measure of postsecondary completion that is not based on certificate or degree completion would be appropriate as long as the definition of completion was made explicit. In this case, completion would be based on whether students fulfilled their own goals as stated upon entering the institution.

When choosing a set of performance measures, states will be sending a message about the priorities of the system.

When choosing a set of performance measures, states will be sending a message to local schools and institutions about the priorities of their system. Therefore, a state should examine the implications of each possible method of measuring an outcome. For example, the choice of a test that is currently administered to all students, vocational and nonvocational alike, could promote the coordination of learning objectives in vocational education with those in all school programs. Ultimately, the choice of this test could represent the beginning of a statewide accountability system for all school performance. Similarly, by choosing to use a definition of high school graduation or postsecondary completion that is already being used by the academic education system, a state could facilitate coordination between vocational and nonvocational programs. However, in some cases, a state may find that existing practices do not meet the needs of the vocational education system.

A state should pay attention to the overall implications of a set of measures as well as the implications of each individual measure. In particular, a state may choose to use more than one measure for each outcome. In order to measure gains in basic and more advanced academic skills, a state may want to use all of the above-mentioned methods, including a test of basic academic skills, an examination of course completion in an advanced academic area, and an alternative assessment such as a writing sample or comprehensive project. If a state were to use only a standardized test of academic skills, it might place too great an emphasis on knowledge of grammar and vocabulary to the exclusion of more complex reasoning abilities, such as the ability to develop an argument. On the other hand, by including a writing sample or comprehensive project, the state could place equal emphasis on the types of skills assessed by standardized tests and those tested in more authentic situations.[2] Furthermore, by examining course completion, the state could spur efforts to develop a coherent sequence of courses that included both academic and vocational competencies.

In addition to ensuring that the choice of measures reflects the goals and targeting priorities of vocational education at both the secondary and postsecondary levels, states should use several additional criteria when evaluating each possible method of measuring an outcome.

- *Availability*. Is the measure currently in use? Which testing instruments and definitions, for example, of program completion and school retention, do schools and institutions currently use? Which measures are used statewide? Which are used more broadly than others? What data already exist for other purposes? How can states be creative in thinking about existing data sources?

- *Comparability*. If a measure is already in use, how comparable is its use across schools and institutions? For example, do all schools or institutions use the same timing of a statewide test? How much variability is there in a current definition?

[2]See Chapter 7 for a complete discussion of the advantages and disadvantages of different types of assessment instruments.

- *Reliability*. If a measure is already in use, how reliable and accurate are the data? For example, what is the response rate on student follow-up or employer satisfaction questionnaires? Are tests administered to all students, to a random sample of students, or to a selected group of students?

- *Timeliness*. If a measure is already in use, how quickly does it produce information?

- *Aggregation*. If a measure is already in use, can the data be attributed to individual students? For example, do schools or institutions currently calculate an average test score or other measure for students, or can scores be attributed to students so that individual gains can be measured?

- *Automation*. How widespread is the use of automated databases across the state? What data currently exist in these databases? What is the capacity of state-level automation?

- *Ease of measuring*. If the measure is not currently in use, how easy would it be for schools and institutions to implement it? How much development work would need to be done before the measure could be widely implemented? How easy would it be to modify existing practices?

- *Cost*. What is the cost of implementing a new measure or of modifying a measure that is already in use? How costly would it be to make an existing measure more comparable, reliable, or timely? Does the measure in question require that the state develop strict data collection guidelines to ensure comparability and reliability? How costly would it be to develop new, alternative assessments of student performance?

States should aim for a set of performance measures that falls somewhere between an ideal set of measures that addresses all of a state's concerns and already existing data collection methods that entail little additional cost. Clearly, some outcomes, such as general occupational competency (all aspects of the industry), as well as other areas of concern in the Perkins Act, such as coherent sequences of courses and integration of academic and vocational education, cannot easily be measured with current methods. The development of appropriate measures will require money, time, and the coordinated effort of state and local educators and other affected parties.

States should aim for a set of performance measures that falls somewhere between the ideal and already existing data collection methods.

However, until there is an ideal system, performance measures can certainly be phased in over time. For example, if statewide assessments of occupational competencies do not currently exist, states could begin in 1992-1993 by allowing local programs to define their own criteria and then have them report to the state on the success of their vocational students according to those criteria. In the meantime, a state-level task force could be working on developing a core set of competencies for each occupational program (i.e., at the postsecondary level) or occupational cluster (i.e., at the secondary level). Furthermore, the state-level task force could develop the core competencies in stages, beginning with the most popular vocational program areas for implementation in 1993-1994 or 1994-1995, and then add smaller programs

at a later date. Although a measure of occupational competency that depends on locally defined criteria may cause some problems for data comparability and reliability, an interim measure will still be useful by allowing schools and institutions to identify their own weaker and stronger programs.

Step 4: Specify the Chosen Measures

After choosing the method(s) to be used to measure each outcome, the state needs to define those measures as specifically as possible. In a highly decentralized system where there is a strong tradition of local control, states may have to define their performance measures in very general terms. However, in a highly centralized system, states have a choice about how specific their measures should be. As discussed in Chapter 1, the process of developing performance measures requires that states decide how much emphasis they want to give to comparability of data at the expense of local autonomy. The more specific the performance measure is, the more consistent the reporting will likely be. A clearly defined performance measure should answer the following questions:

- In what form should the measure be reported?
- To whom should the measure apply? (To whom should the assessment be administered?)
- What measurement method or assessment instrument should be used?
- What time frame should be used? (Under what conditions should the assessment be administered?)

The following are examples of highly specific performance measures that answer the above questions. These measures cover only vocational students, as defined by the state in Chapter 2.

Measures of gains in basic and advanced academic skills at the secondary level.

- The number of normal curve equivalents (NCEs) that a senior gained on the _____ test of reading since taking the test as a sophomore. [what form? NCEs; who? seniors; what method? _____ reading test; what time frame? seniors as sophomores]
- The number of normal curve equivalents (NCEs) that a junior gained on the _____ posttest of math since taking the pretest at the beginning of junior year. [what form? NCEs; who? juniors; what method? _____ pre- and posttest; what time frame? since beginning of junior year]
- Completion of at least one advanced math course by the end of senior year for a student who had not completed any advanced math courses by the end of sophomore year. [what form? whether completed; who? seniors; what method? advanced math course; what time frame? seniors who had not completed as sophomores]

- Successful completion of the posttest of a writing assessment for a freshman who had not successfully completed the pretest at the beginning of the freshman year. [what form? whether completed successfully; who? freshman; what method? pre- and posttest of a writing assessment; what time frame? since beginning of freshman year]

Measures of gains in basic and advanced academic skills at the postsecondary level.

- The number of levels that a student gained on the _____ test of adult basic education for a student who scored poorly upon entering the institution. [what form? number of levels; who? students who scored poorly; what method? _____ test of adult basic education; what time frame? since entering the institution]

- Completion of all general education requirements for a student who had completed all the required courses in a vocational program area in the previous year. [what form? completion; who? students who completed vocational program; what method? general education requirements; what time frame? since previous year]

- Completion of all academic courses that are part of a vocational program area for a student who had completed all the vocational courses in that area in the previous year. [what form? completion; who? students who completed vocational courses in a program area; what method? all academic courses in a vocational program area; what time frame? since previous year]

- Successful completion of the posttest of a writing assessment for a student who had not successfully completed the pretest upon entering the institution. [what form? successful completion; who? students who did not successfully complete; what method? pre- and posttest of a writing assessment; what time frame? since entering the institution]

In addition to answering the four necessary questions, states could include even more detail in their chosen measures. For example, states could specify whether the _____ test of basic reading skills should be given in the Spring or Fall of a student's senior year. However, this level of specificity may conflict with existing local practices. If most schools already administer tests primarily in the Fall, it may be disruptive to require reporting in the spring. Moreover, if academic tests are administered in the Fall to all students by the academic educational division of the state, it may be unworkable to administer the test in the spring just to vocational students. Consequently, when working out the details of performance measures, states should elicit the help and guidance of all affected parties through the Committee of Practitioners.

It is possible for states to define their performance measures in even more specific terms. In order to communicate requirements clearly to local schools and institutions, states can describe their performance measures in terms of an equation, with all relevant variables defined. For example, the first measure of gains in basic academic skills above can be rewritten as follows:

In order to communicate requirements clearly, states can describe their performance measures in terms of an equation.

(NCE level in reading as a senior$_i$ - NCE level in reading as a sophomore$_i$) where i=1 to N,

and

N=total number of vocational students (as defined by the state) taking the _____ test of reading in both the sophomore and senior years

The advantage of defining measures in an equation form demonstrates to the local school or institution precisely how to calculate the measure, that is, what to use as the numerator and what to use as the denominator. This will be particularly helpful when a state holds schools and institutions responsible for aggregating their own data.

Step 5: Translate Student Measures into Program Measures

The above measures are written to measure *student* performance, for instance, the NCE score gain for a student on a reading test or completion of the academic courses in a student's vocational program. However, as explained in Chapter 2, schools and institutions are being held accountable for the performance of their vocational *programs*, or more precisely, for the performance of the students *in their vocational programs*. Consequently, schools and institutions will be evaluated on an aggregate measure of student performance in each relevant program area.

In order to facilitate the evaluation of programs, rather than of individual students, states must translate the above student measures into program measures. Once again, these program measures pertain only to vocational students, as defined by the state in Chapter 2.

Measures of gains in basic and advanced academic skills at the secondary level.

- The average number of normal curve equivalents (NCEs) that seniors gained on the _____ test of reading since taking the test as a sophomore.

- The average number of normal curve equivalents (NCEs) that juniors gained on the _____ posttest of math since taking the pretest at the beginning of the junior year.

- Rate of completion of (or the percentage of students completing) at least one advanced math course by the end of senior year for students who had not completed any advanced math courses by the end of the sophomore year.

- Rate of successful completion of (or the percentage of students completing) the posttest of a writing assessment for freshmen who had not successfully completed the pretest at the beginning of the freshman year.

Measures of gains in basic and advanced academic skills at the postsecondary level.

- The average number of levels that students gained on the _____ test of adult basic education for students who scored poorly upon entering the institution.

- Rate of completion of (or the percentage of students completing) all general education requirements for students who had completed all courses in a vocational program area in the previous year.

- Rate of completion of (or the percentage of students completing) all academic courses that are part of a vocational program area for students who had completed all the vocational courses in that area in the previous year.

- Rate of successful completion of (or the percentage of students completing) the posttest of a writing assessment for students who had not successfully completed the pretest upon entering the institution.

In addition, the previous equation should be respecified as follows to apply to programs:

$$\frac{\text{sum (NCE level in reading as a senior}_i \text{ - NCE level in reading as a sophomore}_i)}{N}$$

where i=1 to N,

and

N=total number of vocational students taking the _____ test in both the sophomore and senior years

In contrast with the first equation, this one produces an average test score gain for all vocational students in a given program area who took the two tests.

Furthermore, several additional measures are possible at the program level. Specifically, average performance can be alternatively written as a percentage or rate.

Measures of gains in basic and advanced academic skills at the secondary level.

- The percentage of seniors gaining **X** normal curve equivalents (NCEs) on the _____ test of reading since taking the test as sophomores.

- The percentage of juniors gaining **X** normal curve equivalents (NCEs) on the _____ posttest of math since taking the pretest at the beginning of the junior year.

Measures of gains in basic and advanced academic skills at the postsecondary level.

- The percentage of students gaining **X** levels on the _____ test of adult basic education for students who scored poorly upon entering the institution.

However, it should be noted that this form of program measure requires including a performance standard. The level of gain, **X**, that is being measured is an implicit standard. The decision to measure "the percentage of students gaining **X**" should not be an arbitrary decision, rather it should be based on a reasonable expectation that students should, or at least could, gain **X** under the given conditions.

All performance measures include some form of implicit standard.

This mixture of performance measures and standards should not be alarming. In fact, all measures include some form of implicit standard. For example, successful completion of a writing assessment, completion of all general education requirements, and completion of all academic courses that are part of a vocational program contain implicit standards. Why not measure, instead, *distinguished* completion of a writing assessment, or completion of *half* of all general education requirements, or completion *with a grade of B or higher* of all academic courses that are part of a vocational program? All measures include an implicit standard because they are based on student outcomes which, in turn, are based on an aspect of student performance.

Performance measures that include implicit standards are appropriate and even unavoidable. The process of defining measures described in this chapter can be followed without first setting those implicit standards. However, before the measures can be implemented, states will need to specify any implicit standards according to the procedure covered in Chapter 4. In the above examples of program measures, states would need to set the **X** levels according to a rational, rather than arbitrary, procedure. In addition, states would need to set any other standards that were buried within the measures. For example, the secondary program measure for advanced academic course completion requires that states set a standard for what constitutes an "advanced math course." Similarly, the postsecondary program measure for gains in reading test scores requires that states set a standard for what constitutes "scored poorly."

Measures of gains in basic and advanced academic skills at the secondary level.

- Rate of completion of (or the percentage of students completing) at least one "advanced math course" by the end of senior year for students who had not completed any "advanced math courses" by the end of the sophomore year.

Measures of gains in basic and advanced academic skills at the postsecondary level.

- The average number of levels that students gained on the _____ test of adult basic education for students who "scored poorly" upon entering the institution.

After consulting with industry and other sources of information covered in Chapter 4, a state might modify the first measure above to read "completion of Algebra I or higher" or "completion of Algebra II or higher." Similarly, a state might modify the second measure to read "scored below X" and then define X. In any case, the appropriate standards will eventually have to be embedded in the relevant measures.

Step #6: Report Program Measures According to the State's Definition of Vocational Programs

The above program measures do not specify which vocational programs should be reported. They could be rewritten with the added phrase "for students in a given vocational program area" as follows:

Measures of gains in basic and advanced academic skills at the secondary level.

- The average number of normal curve equivalents (NCEs) that seniors gained on the _____ test of reading since taking the test as a sophomore *for students in a given vocational program area.*

Measures of gains in basic and advanced academic skills at the postsecondary level.

- The average number of levels that students gained on the _____ test of adult basic education for students who scored poorly upon entering the institution *for students in a given vocational program area.*

As a result, schools or institutions would report the above performance measures for each vocational program area, as defined by the state. If a state defined its vocational programs in terms of the entire vocational program of a school or institution, then the local recipient would be required to report a single statistic. If, on the other hand, a state defined its vocational programs as occupational clusters or detailed occupational programs, then the school or institution would be required to report multiple statistics on each performance measure. Using the above examples, the school or institution would report the average test score gain for students in each occupational cluster or program in the school or institution.

Depending on the state's definition of vocational programs, a school or institution may be required to report multiple statistics on each performance measure.

Examples of Performance Measures

The tables on the following pages offer examples of learning, labor market, and accessibility measures. These measures are generally specified at the program level. It should be cautioned that this is not an exhaustive list and is not meant to take the place of state efforts to develop relevant and appropriate measures.

Table 3.4
Gains in or Attainment of Basic Academic Skills

Examples of Performance Measures at the Secondary Level

Standardized Test Score Gains

The average number of points that vocational seniors gained on the _____ test of reading/language/math/science/social studies skills since they took the test as sophomores.

The average number of points that vocational students gained between the pre- and posttests of the _____ test of reading/language/math/science/social studies skills out of the total number of vocational students taking both tests.

The percentage of vocational students gaining **X** or more points between the pre- and posttests of the _____ test of reading/language/math/science/social studies skills out of the total number of vocational students scoring below proficiency on the pretest.

Standardized Test Score Attainment

The percentage of vocational seniors scoring **X** or above on the _____ test of reading/language/math/science/social studies skills out of the total number of vocational seniors taking the test.

The percentage of vocational seniors scoring **X** or above on a retest of the _____ test of reading/language/math/science/social studies skills out of the total number of vocational seniors who scored below **X** on the initial test.

The percentage of vocational sophomores scoring **X** or above on the _____ writing assessment out of the total number of vocational sophomores participating in the assessment.

Academic GPA

The percentage of vocational students maintaining a 2.0 GPA in all academic courses out of the total number of vocational students.

The average gain in GPA for academic courses only from the end of the first semester to the end of the second semester for all vocational students.

Alternative Assessment

The percentage of vocational seniors completing a portfolio requirement with a passing or proficient grade out of the total number of vocational seniors participating in the portfolio assessment.

Table 3.4
Gains in or Attainment of Basic Academic Skills—Continued

Examples of Performance Measures at the Postsecondary Level

Test Score Gains

The average number of points that second-semester vocational students gained on the _____ test of adult basic education skills since taking the test upon college entrance for all vocational students who did not achieve a minimum score of **X** upon entrance.

The average number of points that vocational students gained between the pre- and posttests of the _____ test of reading/language/math skills out of the total number of vocational students enrolled at the institution who scored below **X** on the same test during high school.

Test Score Attainment

The percentage of vocational students scoring **X** or above on the _____ test of adult basic education skills out of the total number of vocational students taking the test.

The percentage of vocational students scoring **X** or above on a retest of the _____ test of reading/language/math skills out of the total number of vocational students enrolled at the institution who scored below **X** on the same test during high school.

Course Completion

The percentage of vocational students successfully completing developmental (remedial) courses out of the total number of vocational students who were referred to developmental courses.

The average number of developmental (remedial) credits *earned* by vocational students out of the total number of developmental (remedial) credits *attempted* by vocational students during the reporting period.

GED Completion

The percentage of vocational students entering college without a high school diploma who successfully obtained a GED or other equivalent within one year.

Alternative Assessment

The percentage of second-year vocational students completing with a passing or proficient grade a written paper on a subject in their vocational area out of the total number of vocational students who participated in the writing assignment.

Table 3.5
Gains in or Attainment of More Advanced Academic Skills

Examples of Performance Measures at the Secondary Level

Test Score Attainment

The percentage of vocational seniors attaining X or higher on an Advanced Placement exam or X or higher on a College Board Achievement Test out of the total number of vocational seniors taking either test.

Course Completion

The percentage of vocational seniors completing Algebra II sometime during their high school career out of the total number of vocational seniors. [An advanced math course could be defined differently, for instance, as Algebra I or Geometry instead of Algebra II.]

The percentage of vocational seniors completing Algebra I by the end of their senior year out of the total number of vocational seniors who had not completed Algebra I at the beginning of their senior year.

The percentage of vocational seniors completing chemistry or physics sometime during their high school career out of the total number of vocational seniors. [An advanced science course could also be defined differently.]

The average number of credits earned by vocational completers in advanced academic courses. [Advanced academic courses would have to be defined. For instance, one state includes the following courses: Algebra I & II, Geometry, Trigonometry, Calculus, Applied Math (CORD) I & II, Biology I, Advanced Biology, Physics, Chemistry I, Advanced Chemistry, Principles of Technology (CORD, AIT) I & II, Applied Biology/Chemistry (CORD), and AP English.]

Academic GPA

The percentage of vocational students maintaining a 2.0 GPA in all advanced academic courses out of the total number of vocational students who took at least one advanced academic course.
The average gain in GPA for advanced academic courses only from the end of the first semester to the end of the second semester for all vocational students.

Alternative Assessment

The percentage of vocational seniors completing with a passing or proficient grade an extensive research project on a topic in their vocational area, and that was conducted in conjunction with an academic course, out of the total number of vocational seniors.

Table 3.5
Gains in or Attainment of More Advanced Academic Skills—Continued

Examples of Performance Measures at the Postsecondary Level

Test Score Attainment

The percentage of second-year vocational students attaining X or higher on an Advanced Placement exam or X or higher on a College Board Achievement Test out of the total number of second-year vocational students taking either test.

Course Completion

The percentage of vocational students who completed all academic course requirements associated with their specific vocational program area out of the total number of vocational students who completed the vocational course requirements associated with their specific vocational program area.

The percentage of vocational students who completed all general education requirements out of the total number of vocational students who completed the vocational course requirements associated with their specific vocational program area.

The percentage of vocational students who completed selected general education courses out of the total number of vocational students who completed the vocational course requirements associated with their specific vocational program area.

The percentage of vocational completers who successfully completed a college-level algebra course (academic or applied).

The percentage of vocational completers who successfully completed a laboratory science course.

The average number of advanced academic credit hours *earned* by vocational students out of the total number of advanced academic credit hours *attempted* by vocational students during the reporting period. [Advanced academic courses would have to be defined.]

Academic GPA

The percentage of vocational students who maintained a GPA of 2.0 or higher in all academic courses taken that semester/year out of the total number of vocational students enrolled that semester/year.

The percentage of vocational completers who maintained a GPA of 3.0 or higher in all academic courses out of the total number of vocational completers.

Alternative Assessment

The percentage of vocational completers successfully completing an extensive research project on a topic in their vocational area and that was conducted in conjunction with an academic course or instructor out of the total number of vocational completers.

Table 3.6
Gains in or Attainment of Specific Occupational Competency

Examples of Performance Measures at the Secondary Level

Competency Tests

The percentage of vocational seniors completing the _____ test of occupational competencies with a passing or proficient score out of the total number of vocational seniors taking the test.

The percentage of vocational seniors passing a retest of the _____ test of occupational competencies out of the total number of vocational seniors who did not pass the initial test.

Competency Checklists

The percentage of vocational program completers attaining X percent of the competency objectives in their vocational program area out of the total number vocational program completers.

Vocational GPA

The percentage of vocational students maintaining a GPA of 2.0 or higher in all vocational courses taken that semester/year out of the total number of vocational students taking vocational courses that semester/year.

The percentage of vocational completers maintaining a GPA of 3.0 or higher in all vocational courses out of the total number of vocational completers.

Eligibility for Co-Op

The percentage of vocational seniors who are eligible to participate in co-op out of the total number of vocational seniors (i.e., the percentage of vocational seniors demonstrating a minimum vocational competency level).

Program Completion

The percentage of vocational seniors receiving certification in their vocational program area out of the total number of vocational seniors (i.e., the percentage of vocational seniors demonstrating proficiency in the vocational competencies in their program area).

Performance-Based Assessment

The percentage of vocational seniors who complete with a passing or proficient grade a project in their vocational program area that results in a physical product or measurable service out of the total number of vocational seniors who participate in the project.

Licensing Exams

The percentage of vocational completers who successfully pass the state, national licensing, or certification exam or otherwise fulfill the state, national licensing, or certification requirements in their vocational field out of the total number of vocational completers who are eligible to be licensed or certified.

Employer Evaluations

The percentage of vocational students participating in co-op in a job related to their vocational program area who received a satisfactory evaluation of their occupational-specific skills from their co-op employer.

The percentage of vocational completers employed in a job related to their vocational program area who received a satisfactory evaluation of their occupational-specific skills from their employer.

Table 3.6
Gains in or Attainment of Specific Occupational Competency—Continued

Examples of Performance Measures at the Postsecondary Level

Competency Tests

The percentage of vocational students completing a test of occupational competencies with a passing or proficient score out of the total number of vocational students taking the test.

Competency Checklists

The percentage of vocational completers attaining X percent of vocational competencies for their program out of the total number of vocational program completers.

Vocational GPA

The percentage of vocational students who maintained a GPA of 2.0 or higher in all required courses in their vocational program taken that semester/year out of the total number of vocational students enrolled that semester/year.

The percentage of vocational completers who maintained a GPA of 3.0 or higher in all required courses in their vocational program out of the total number of vocational completers.

Course Completion

The percentage of vocational students completing all core courses or other specified sequences of courses in a vocational program area out of the total number of vocational students.

Eligibility for Clinical Practice

The percentage of second-year vocational students eligible for clinical practice out of the total number of second-year vocational students (i.e., the percentage of second-year vocational students demonstrating a minimum vocational competency level).

Performance-Based Assessment

The percentage of second-year vocational students who complete with a passing or proficient grade a work simulation test out of the total number of second-year vocational students taking the test.

Licensing Exams

The percentage of vocational completers who successfully pass the state, national licensing, or certification exam or otherwise fulfill the state, national licensing, or certification requirements in their vocational field out of the total number of vocational completers who are eligible to be licensed or certified.

Employer Evaluations

The percentage of vocational students participating in clinical practice in a job related to their vocational program who received a satisfactory evaluation of their occupational-specific skills from their clinical supervisor.

The percentage of vocational completers employed in a job related to their vocational program area who received a satisfactory evaluation of their occupational-specific skills from their employer.

Table 3.7
Attainment or Enhancement of General Occupational Skills*

Examples of Performance Measures at Either the Secondary or Postsecondary Levels

Course Completion

The percentage of vocational completers who successfully completed a course covering one or more of the following aspects of the industry that the student was preparing to enter out of the total number of vocational completers: planning, management, finance, technical and production skills, underlying principles of technology, labor and community issues, health and safety, and environment issues. For example, a course offered in the vocational program area may cover one or more of these aspects as related to the specific industry that the student is preparing to enter.

Tests of Understanding

The percentage of vocational completers who completed a test (multiple choice, short answer, and/or essay) with a passing or proficient grade that assesses the student's understanding of one or more aspects of the industry that the student is preparing to enter.

Alternative Assessments

The percentage of vocational students who successfully completed a written paper relating a course in one of the relevant aspects of industry to the specific industry that the student was preparing to enter out of the total number of vocational students taking a course in one of the relevant aspects of industry. For example, a student who is taking a finance course may write a paper on a topic that relates what the student is learning in finance to the specific industry that the student plans to enter.

The percentage of vocational completers who successfully completed a performance-based assessment that integrates experience in and understanding of all aspects of the industry that a student is preparing to enter out of the total number of vocational completers. For example, a student may participate in a senior or degree project that involves working in the community, such as participating in the construction of affordable housing or in the establishment of a neighborhood clinic, and that requires the student to confront the range of issues faced by people working in the industry.

* The Notice of Proposed Rulemaking § 400.4(b) defines *general occupational skills* as meaning "experience in, and understanding of, all aspects of an industry," and further defines *all aspects of an industry* as including "planning, management, finance, technical and production skills, underlying principles of technology, labor and community issues, health and safety, and environment issues related to the industry" that a student is preparing to enter.

Table 3.8
Attainment or Enhancement of Employability Skills (Workplace Basics)*

Examples of Performance Measures at Either the Secondary or Postsecondary Levels

Standardized Tests

The percentage of vocational students scoring X percent on the _____ test of employability skills out of the total number of vocational students taking the test.

Skills Checklists

The percentage of vocational completers attaining X percent on a checklist of employability tasks out of the total number of vocational completers.

Participation in Guidance Activities

The percentage of vocational students satisfactorily participating in career exploration activities and developing a personal work and further education plan through the guidance department out of the total number of vocational seniors.

The percentage of vocational students satisfactorily completing a workshop on work readiness out of the total number of vocational students.

Alternative Assessments

The percentage of vocational completers demonstrating mastery of specified employability skills through a series of observations and/or role plays out of the total number of vocational completers. The level of performance that satisfies mastery will need to be explicitly defined.

The percentage of vocational students completing with a passing or proficient grade a work-readiness portfolio (including such elements as a resume, writing sample, and letter of recommendation) out of the total number of vocational students participating in the portfolio assessment.

* States have included the following among their workplace skills: managing resources such as time and material, participating as a team member, using computers to process information, decision making and problem solving, communication skills such as listening and speaking, learning strategies, exhibiting a proper work ethic, creative thinking, taking responsibility, appearance, attitude, ability to lead and to follow, ability to manage personal and family life, ability to identify and set education and employment goals, and job search skills.

BEST COPY AVAILABLE

Table 3.9
Retention in School

Examples of Performance Measures at the Secondary Level

Graduation Rate

The percentage of vocational completers who graduated from high school out of the total number of vocational completers in the graduating class.

The percentage of students enrolled in but not completing vocational programs who graduated from high school out of the total number of vocational enrollees in the graduating class.

The percentage of students enrolled in but not completing vocational programs who dropped out of school but who completed requirements for a GED or high school diploma within the last five years out of the total number of vocational enrollees in the graduating class of five years ago.

Retention Rate

The percentage of vocational students still enrolled in school at the end of the academic year out of the total number of vocational students who were enrolled in school at the beginning of the academic year, not including transfers in and out.

Attendance Rate

The percentage of days that vocational students were in attendance out of the total number of vocational student days (ADA).

The percentage of days that vocational students were in attendance and were neither tardy nor dismissed early out of the total number of vocational student days.

The percentage of class periods for which vocational students were present out of the total number of scheduled vocational student class periods.

Table 3.9
Retention in School—Continued

Examples of Performance Measures at the Postsecondary Level

GED/Diploma Completion

The percentage of students who were vocational students in high school and who entered the postsecondary institution X years/semesters ago without a high school diploma or equivalent who completed requirements for a GED or high school diploma within the last X years/semesters.

The percentage of students who entered the postsecondary institution X years/semesters ago without a high school diploma or equivalent and who stated that they had vocational goals at time of entrance who completed requirements for a GED or high school diploma within the last X years/semesters.

Certificate or Degree Completion

The percentage of vocational students who entered the institution X years/semesters ago who completed a certificate or degree within the last X years/semesters.

The percentage of students who completed the requirements of their vocational program X years/semesters ago and who were awarded a degree within the last X years/semesters.

Goal Completion

The percentage of students who stated that they had vocational goals at time of entrance to the institution X years/semesters ago who completed those goals within the last X years/semesters.

Retention Rate

The percentage of vocational students who were enrolled during the Fall semester and who returned during the spring semester.

The percentage of vocational students who were enrolled sometime during the previous academic year and who were enrolled again sometime during the current academic year.

The percentage of vocational students who were enrolled sometime during the current academic year and who earned at least X credits during the current academic year.

Table 3.10
Retention in Vocational Programs

Examples of Performance Measures at the Secondary Level

Program Completion

The percentage of students in the graduating class who completed a vocational program out of the total number of students in the graduating class who were enrolled in a vocational program.

Completion of a Coherent Sequence of Courses

The percentage of students in the graduating class who completed a coherent sequence of courses leading to both academic and vocational competencies out of the total number of students in the graduating class who completed the first course in a coherent sequence.

Course Completion

The percentage of students who *completed* at least one vocational course during the current academic year out of the total number of students who were *enrolled* in at least one vocational course during the current academic year.

The percentage of credits *earned* in vocational courses by students during the current academic year out of the total number of vocational credits *attempted* during the current academic year.

Participation in Tech Prep

The percentage of students completing the secondary portion of the tech-prep program out of the total number of students beginning the secondary portion of the tech-prep program.

The percentage of students enrolling in the postsecondary portion of the tech-prep program out of the total number of students completing the secondary portion of the tech-prep program.

Table 3.10
Retention in Vocational Programs—Continued

Examples of Performance Measures at the Postsecondary Level

Program Completion

The number of first-time, full-time enrollees completing vocational programs out of the total number of first-time, full-time students declaring an intent to complete a vocational program.

The number of students completing vocational programs out of the total number of students enrolling in a vocational program.

Completion of a Coherent Sequence of Courses

The percentage of students completing a planned sequence of vocational and academic courses out of the total number of students taking vocational courses.

Course Completion

The percentage of credits *earned* in vocational courses by students during the current academic year out of the total number of vocational credits *attempted* during the current academic year.

Participation in Tech Prep

The percentage of students completing the postsecondary portion of the tech-prep program out of the total number of enrolled students who completed the secondary portion of the tech-prep program.

Table 3.11
Placement and Retention in the Labor Market

Examples of Performance Measures at Both the Secondary and Postsecondary Levels

Placement Rate

Percentage of program completers who are either employed, employed in a job related to training, employed in the military or other service such as Peace Corps, or enrolled in further education or training. Each type of placement could be reported separately.

Employment Rate

Percentage of program completers who are employed in a given calendar quarter.

Time to Employment

The average period of time it took program completers to be employed in their first job.

Unemployment Rate

Percentage of program completers who are unemployed at a given point in time. "Unemployed" means not in the labor market *and* looking for work.

The average period of unemployment for program completers since exiting the program.

Job Retention Rate

The average period of time program completers were employed in their first job.

Percentage of program completers who were employed in calendar quarter "i + 1" compared with the percentage of program completers who were employed in calendar quarter "i."

Table 3.12
Earnings

Examples of Performance Measures at Both the Secondary and Postsecondary Levels

Average Wage

The average entry-level wage earned by vocational program completers in their first job after exiting the program.

The average wage earned by vocational program completers at a given point in time after exiting the program.

Average quarterly earnings of vocational program completers for a given calendar quarter.

Rate of Earnings Increase

The rate at which the entry-level wage earned by vocational program completers increased over a given period of time after exiting the program.

The rate at which quarterly earnings increased for vocational program completers.

Table 3.13
Employer or Employee Satisfaction

Examples of Performance Measures at Both the Secondary and Postsecondary Levels

Employer Satisfaction

The percentage of employers rating vocational program completers at a satisfactory level or higher as measured by an employer survey.

Employee Satisfaction

The percentage of vocational program completers rating their first job related to training at a satisfactory level or higher as measured by an employee survey.

Table 3.14
Access to and Success in Vocational Education for Special Populations

Examples of Performance Measures at the Secondary Level

Enrollment Rates

Percentage of students who are members of special populations enrolled in vocational education programs compared with the percentage of students who are members of special populations in the school population at large.

Completion Rates

Percentage of vocational program completers who are members of special populations compared with the percentage of students who are members of special populations in the school population at large.

Percentage of vocational program completers who graduated and who are members of special populations compared with the percentage of graduating students who are members of special populations.

Performance on Other Measures

The performance (percentage or average score) of members of special populations on any previous performance measures compared with the performance of the regular student population on the same measure.

Table 3.14
Access to and Success in Vocational Education for Special Populations—Continued

Examples of Performance Measures at the Postsecondary Level

Enrollment Rates

Percentage of vocational-technical enrollees who are members of special populations compared with the percentage of nonvocational enrollees who are members of special populations.

Percentage of enrollees in each vocational-technical program (reported separately) who are members of special populations compared with the percentage of nonvocational enrollees who are members of special populations.

Where the entire student population of an institution is vocational-technical the percentage of enrollees who are members of special populations compared with the percentage of adults residing in the county (or other geographic area) who are members of special populations.

Completion Rates

Percentage of vocational program completers who are members of special populations compared with the percentage of students who are members of special populations in the school population at large.

Performance on Other Measures

The performance (percentage or average score) of members of special populations on any previous performance measures compared with the performance of the regular student population on the same measure.

4. Performance Standards

After states have defined their performance measures, they are ready to set standards for those measures. This chapter describes several key issues that states need to consider when setting appropriate standards and outlines a process that they can use as a guideline.

No Right or Wrong

The setting of performance standards requires investigation, flexibility, and sound judgment. There is no single right method. By endeavoring to set standards for vocational education programs, vocational educators are participating at the forefront of current efforts to establish standards for all education in the United States. As part of this vanguard, states will find that there are few established rules, few tried and true methods, and few ready-made standards that can be adopted. While this chapter offers some guidelines to follow and suggests some routes to pursue, states should expect the process of setting standards to include a fair amount of trial and error.

States should expect the process of setting standards to include a fair amount of trial and error.

The Role of Performance Standards in the Accountability System

At the end of academic year 1992–1993, schools and institutions will begin to assemble data on the state's performance measures. How should this information be interpreted? While the Perkins Act requires that recipients annually evaluate their vocational education programs, on what basis should they determine that a program is successful or, inversely, that it requires improvement efforts? Performance standards provide schools and institutions, and the state, with a benchmark against which to evaluate programs. The performance of a program in relation to these standards will indicate what course of action, if any, should be taken.

Performance in the "Real World" As a Guiding Principle

When setting standards, states should make sure that the chosen level of performance has been demonstrated as necessary for success in either the workplace or in further education and training. States should try to avoid choosing standards on an arbitrary basis that cannot be validated in a real world context. Otherwise, standards will be placed under scrutiny, and the integrity of state accountability systems will be challenged. As a guiding principle, states should constantly ask whether there is sufficient evidence that achievement of the chosen level of competence is necessary to perform successfully in either the workplace or in further education and training.

Standards must have validity in the workplace or in further education or training.

Piecing Together the Picture of Necessary Performance

It is not a simple task to determine the level of performance on each measure that is necessary for success in the workplace or in further education and training. Although states can ask employers, postsecondary institutions, and institutions of higher learning what they believe to be the necessary skills, knowledge, and abilities to succeed in their respective spheres, validating that this performance is indeed necessary is probably beyond the scope of what states can be expected to achieve in their accountability systems. Consequently, states are forced to accept indirect indications of necessary levels of performance. For example, states may use the qualifications that an employer requires for an entry-level position, or the entrance requirements at a postsecondary institution, as a proxy for the levels of performance that will be necessary to succeed in those places.

Furthermore, the skills, knowledge, and abilities needed to succeed in a particular place of work or institution may not be the same skills, knowledge, and abilities that vocational students will need throughout their work or school lives. By relying on information gathered locally, or even within a state, states may be short-changing their students. While some vocational students can probably expect to work throughout their lives in a local labor market, many will find themselves affected by national and international markets. Determining what levels of performance are necessary to prepare a student for success in an ever-changing economy is an even more complicated task.

States should elicit input on needed skills from as broad a range of community members as possible.

In the absence of reliable and verifiable information on the level of performance that vocational students will need for success throughout their work and school lives, states must rely on approximations of those necessary skills, knowledge, and abilities. However, several guidelines should be helpful. First, states should elicit input on needed skills from as broad a range of community members as possible, including representatives of industry and higher education, as well as parents and students. Only through participation of the community will states be able to piece together a comprehensive picture of what vocational students may need from their education. Second, while respecting the needs of local and regional labor markets, states should set their sights high and aim to develop a set of world class standards that will serve students well into the future.

Sources of Information on Standards

Where can states look for information on setting appropriate standards? The following sources can provide information that serves as a starting point for discussion. However, final decisions about performance standards should rely on information gathered from a variety of sources as well as the best judgment of those participating in the process. Four strategies to consider are:

- Departure from averages
- Industry standards
- Best practices
- State or national certification or licensing requirements

Departure from Averages

An initial source of information on levels of performance can come from state or national (or even international) averages on standardized tests, or from average performance on the measures collected by the state's system of measures and standards. However, as explained in Chapter 1, there are a number of problems associated with setting standards from average performance. First, averages provide no external validation that a level of performance is desirable. Second, as described in Chapter 7, averages, by definition, divide students or programs into those that are above the average and below, and in effect create winners and losers. Even if a school or institution's performance increases from one year to the next, that recipient may still be ranked the same in comparison with other schools or institutions in the state. The effect of ranking performance based on a normal curve distribution is to create competition between schools and institutions, rather than to challenge a school or institution to improve performance based on external, objective criteria. However, in some cases, average performance on standardized tests may be the only available information.

An initial source of information can come from averages on standardized tests or from average performance on the state's core measures.

While the departure from averages method can help states determine what standards to set with respect to specific measures, we suggest that standards be based on criteria that can be externally validated, in other words, that have been demonstrated as necessary for success in either the workplace or further education or training. As discussed later in this chapter, the departure from averages approach can be a useful management tool for deciding who needs to develop a program improvement plan.

Industry Standards

States should seek the advice of industry as to employer expectations. Unfortunately, few industries have developed standards at the national level, and those that have done so have tended to set standards for the content of programs, such as equipment, topics covered, and instructor qualifications, rather than the expected performance of students after completing those programs. However, by approaching industry with a request for help with setting standards, states may provide a catalyst for industry to begin to examine their skill requirements.

In the meantime, schools and institutions should strengthen their relationship with industry through their technical advisory committees. Instead of focusing on program inputs, such as equipment or instructor qualifications, states should involve these local employers in determining the levels of performance that students should master upon completion of a vocational program. In addition, states may want to try to broaden the local focus of advisory committees to include state or regional level discussions about standards. This may be particularly helpful when developing standards for occupational clusters that cover a range of specific occupations.

States may want to broaden the local focus of advisory committees to include state or regional level discussions about standards.

Best Practices

Best practices refers to standards that have already been adopted by an exemplary institution or program. States can find an indirect indicator of necessary levels of student performance by examining the practices of exemplary institutions or vocational programs in the state, country, or even in the international realm. One of the advantages of relying on existing exemplary practices is that there is no question about whether the level of performance is attainable. Instead, schools and institutions should ask why another program is able to achieve that level of performance. Thus, the exemplary school or institution serves as as laboratory for trying different strategies for vocational education.

State or National Certification or Licensing Requirements

Finally, the levels of performance required on state or national certification or licensing exams provide another source of information on the skills, knowledge, and abilities required to enter certain occupations. However, certification and licensing are currently limited to a relatively small number of occupations. Furthermore, states may have to be creative about extrapolating this information to determine the levels of performance required for an occupational cluster or industry.

Setting Standards Relevant to Chosen Student Measures

Student performance forms the basis for all standard setting.

Standards should first be set for student performance. It is student performance that forms the basis of all standard setting, since standards are concerned with the level of performance required of students to succeed in the workplace and in further education and training. After addressing student performance, standards can then be set for other levels of aggregation such as programs. Consequently, the process of setting standards should begin with the student measures defined in Chapter 3. For example,

Measures of gains in basic and advanced academic skills at the secondary level

- The number of normal curve equivalents (NCEs) that a senior gained on the _____ test of reading since taking the test as a sophomore for all vocational students who took both tests.

  Sample standard: Gain of 10 NCEs.*

*All numerical standards used in this guide are meant only as illustrations and are not intended as suggested or recommended standards.

Measures of gains in basic and advanced academic skills at the postsecondary level

- The number of levels that a student gained on the _____ test of adult basic education for a student who scored poorly upon entering the institution for all vocational students who took both tests.

  Sample standard: Gain of 2 levels.*

In the first case, states would have to set a standard for the number of normal curve equivalents a senior should be expected to gain on the _____ test of reading since taking the test as a sophomore. In the second case, states would have to set a standard for the number of levels a student who scored poorly upon entering the institution should be expected to gain on the _____ test of adult basic education. (In addition, states would have to decide what "scored poorly" means.) When using standardized tests as the method of measuring performance, states have little choice but to base their standards on average performance. However, states have some flexibility in setting standards in this situation. For example, states could choose to set the above standards based on the average performance of all students in the state, rather than just vocational students, or, alternatively, on the national average for all students or for vocational students. In any case, states must set a standard for expected student performance relative to the chosen measure.

The following are a different set of student measures drawn from Chapter 3:

Measures of gains in basic and advanced academic skills at the secondary level

- Completion of at least one advanced math course by the end of senior year for a student who had not completed any advanced math courses by the end of sophomore year for all vocational students.

  Sample standard: Completion of Algebra I or higher.*

Measures of gains in basic and advanced academic skills at the postsecondary level

- Successful completion of the post-test of a writing assessment for a student who had not successfully completed the pretest upon entering the institution for all vocational students.

  Sample standard: Completion of the post-test with a score of B or higher.*

In the first case, states would have to define for what constitutes an "advanced math course" at the secondary level. In the second case, states would have to set the criteria to be used for an evaluation of "successful completion" of a writing assessment at the postsecondary level. In order to determine the level of advanced math that students require, states may

*All numerical standards used in this guide are meant only as illustrations and are not intended as suggested or recommended standards.

survey employers as well as postsecondary institutions. Similarly, to determine what constitutes successful completion of a writing assessment, states may survey employers about the writing skills expected of postsecondary program completers.

States should consider whether different standards should be set for students in different vocational programs.

States should also consider whether different standards should be set for students in different occupational clusters or programs. For example, the level of advanced math that a student requires for success in the workplace might vary for students in welding, health, and drafting programs. Similarly, the level of writing ability required of students in business occupations may differ from the level required of students in auto mechanics. However, states may prefer to set one overall standard for the level of math that students might require throughout their work lives.

Setting Standards Relevant to Chosen Program Measures

After setting standards for student performance, states should set standards for program measures. The following examples of program measures are drawn from Chapter 3:

Measures of gains in basic and advanced academic skills at the secondary level

- The average number of normal curve equivalents (NCEs) that seniors gained on the _____ test of reading since taking the test as sophomores.

  Sample standard: Gain of 10 NCEs.*

Measures of gains in basic and advanced academic skills at the postsecondary level

- The average number of levels that students gained on the _____ test of adult basic education for students who scored poorly upon entering the institution.

  Sample standard: Gain of 2 levels.*

In these two cases, the standards set in the previous section for expected student gains on the relevant tests should be the same as the standards for the average gains made by students in the same programs. This is because of the nature of averages. However, it is also possible that a state might set a single overall standard for all students, and then set separate standards for students in each program area.

Measures of gains in basic and advanced academic skills at the secondary level

- The percentage of seniors gaining **X** normal curve equivalents (NCEs) on the _____ test of reading since taking the test as sophomores.

  Sample standard: 100 percent of students will gain 10 NCEs.*

*All numerical standards used in this guide are meant only as illustrations and are not intended as suggested or recommended standards.

Measures of gains in basic and advanced academic skills at the postsecondary level

- The percentage of students gaining **X** levels on the _____ test of adult basic education for students who scored poorly upon entering the institution.

  Sample standard: 100 percent of students will gain 2 levels.*

In the above two examples, the earlier standards (**X**) set for student gains on the relevant tests should be incorporated into the program measures. Then, a second-level standard should be set for the percentage of students in a program who are expected to achieve that level. If the standard (**X**) accurately describes the gain a student should be expected to attain, then ultimately the percentage of students achieving the gain standard would be one hundred percent. However, the percentage of students in a program in a given year who achieve the world class standard **X** might reasonably be expected to be lower than this level.

Student standards should be incorporated into program measures.

Measures of gains in basic and advanced academic skills at the secondary level

- Rate of completion of (or the percentage of students completing) at least one advanced math course by the end of senior year for students who had not completed any advanced math courses by the end of sophomore year.

  Sample standard: 100 percent of students will complete Algebra I or higher.*

Measures of gains in basic and advanced academic skills at the postsecondary level

- Rate of successful completion of (or the percentage of students completing) the post-test of a writing assessment for students who had not successfully completed the pretest upon entering the institution.

  Sample standard: 100 percent of students will complete the post-test with a score of B or higher.*

In the above two examples, the standards set previously on the associated student measures should be incorporated into the program measures. For example, the definition of an "advanced math course" should be incorporated into the first measure, and the criteria for successful completion of a writing assessment should be incorporated into the second measure. Then, a new second-level standard should be set for the percentage of students achieving those levels of performance. Once again, states may have to rely on average performance to determine what program standard should be set. For example, if fifty percent of vocational students statewide either completed Algebra I or demonstrated competency in the set writing criteria, then states

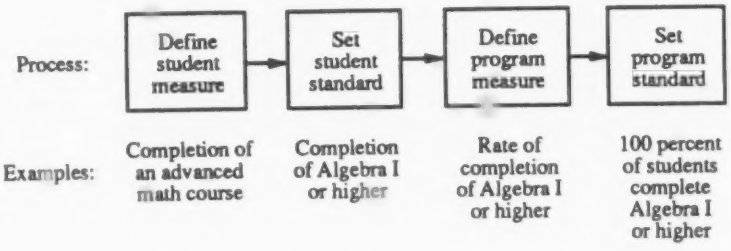
After incorporating student standards into program measures, a second-level standard should be set for the program.

*All numerical standards used in this guide are meant only as illustrations and are not intended as suggested or recommended standards.

may use that average performance as a departure point for setting an appropriate standard on the relevant completion rates.

Figure 4.1 summarizes the standard setting process.

Figure 4.1
Process of Setting Student and Program Standards


Process:
Define student measure
Set student standard
Define program measure
Set program standard
Examples:
Completion of an advanced math course
Completion of Algebra I or higher
Rate of completion of Algebra I or higher
100 percent of students complete Algebra I or higher

General Considerations

Several additional considerations should inform the process of setting standards.

Set Reasonable Levels of Performance that Do Not Force Failure

When setting the level of performance to which students and programs will be held accountable, states should be careful to set a level that represents an ideal state of performance on a measure, and that presents a challenge to vocational programs, but does not push the standard so far that people will necessarily fail. For example, while it might be ideal that one hundred percent of vocational students take Algebra I by their senior year, setting this standard is probably unrealistic. Even if ninety-five percent of students in a program were to take Algebra I, the program would still fall short of the standard. Although a level of one hundred percent might be reasonable in some cases, such as demonstrating competency in brake repair or in administering an injection to a patient, states should probably consider setting a standard at a slightly lower level.

Set Standards for Outcome Measures Only

Although states may find it efficient to include process variables such as program quality factors and other student and institutional characteristics, among their chosen outcomes and measures, states should be cautious about setting standards for process measures. It is one thing to collect supplemental information on programs and quite another to hold schools and institutions accountable for these inputs. An outcome-based system focuses

on how well vocational programs serve students while leaving the specific strategies to local schools and institutions. By setting standards for inputs, states may be jumping to conclusions about what local strategies work. If a state does decide to include process variables among its outcomes and measures and to set standards for those measures, then the state should be confident that those inputs are directly linked to successful outcomes.

Monitor Performance in Relation to Standards and Make Necessary Adjustments

Over time, states should monitor the performance of programs with respect to the chosen standards and adjust those standards as necessary. This is especially important where standards are based on average performance rather than on external objective criteria. As programs improve, standards can be adjusted higher. However, as stated above, states should be careful to avoid the situation where standards are set so high, or are raised so frequently, that program improvement efforts are not recognized and schools and institutions feel doomed to fail.

Triggering the Program Improvement Plan Process

The Perkins Act requires that schools and institutions annually evaluate their vocational education programs based on the system of performance measures and standards. If, after one year, local recipients do not make "substantial progress" in meeting the standards and measures, they are required to develop a program improvement plan. However, the Act does not define "substantial progress" or indicate how the improvement process should be set in motion. Consequently, states must decide what mechanism will be used to trigger the program improvement plan process.

States must decide what mechanism will be used to trigger the program improvement process.

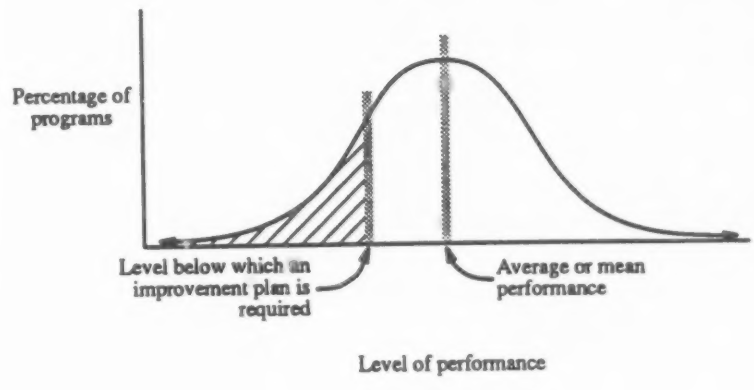
As a result of the steps taken in Chapters 2 and 3, and in the earlier part of this chapter, states will soon have a substantial amount of data on program performance available. In addition, states will know how that performance compares with the state standards. Two main scenarios are possible: (1) where standards were set according to average performance; and (2) where standards were set according to external objective criteria. In the first case, somewhere close to one-half of all programs should fall below the state average. In the second case, it is initially unknown how many programs will fall below the standard. In each case, how should the state determine which programs should be required to develop a program improvement plan?

Triggering the Program Improvement Process in the First Year of the System

Where standards are set according to average performance, states could require that all programs performing below the state average should develop a program improvement plan; however, this would keep a lot of programs very busy on a continual basis. Alternatively, states could require only those programs performing one or more standard deviations below the average to

develop plans. This should theoretically result in about sixteen percent of programs being required to develop plans. Moreover, states could simply require the lowest quintile (twenty percent) or quartile (twenty-five percent) to develop plans. Where standards are set according to external objective criteria, states could require programs that perform in the lowest quintile or quartile in relation to an externally validated standard, or some other low-performing classification, to develop a program improvement plan. In any case, these methods should probably only be implemented for the first round of evaluations. After the first year, states would probably want to go to a system that recognized individual program progress, rather than always requiring the same lowest performing group to develop a plan every year (Figure 4.2).

Figure 4.2
Triggering the Program Improvement Plan in the First Year


Percentage of programs
Level below which an improvement plan is required
Average or mean performance
Level of performance

Setting Annual Performance Targets

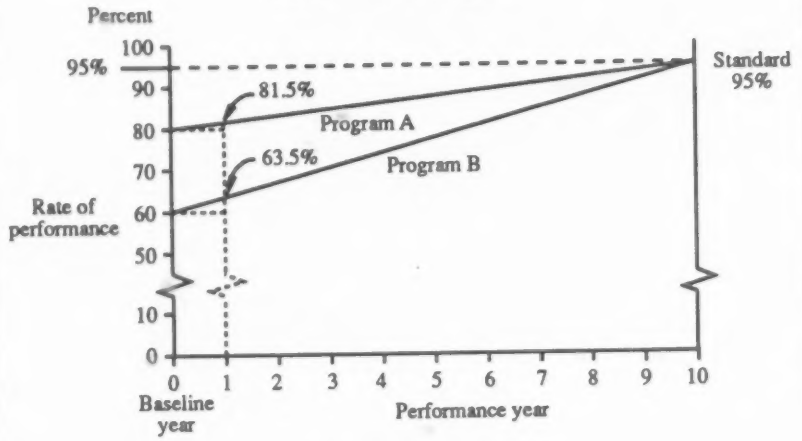
After the first year, states will probably want to shift to a system that recognizes annual program progress.

After the first year of collecting data on program performance, states will probably want to shift to a system that recognizes annual program progress. Such a system would set annual performance targets against which programs would be held accountable, rather than keeping program performance tied to the overall standard.

States should attempt to set annual performance targets that keep in mind the overall performance standard and that are fair in their treatment of differing programs. One such system bases performance targets on the percentage decrease between the state standard and the previous year's actual performance. While all programs would be held accountable for the same percentage decrease, the absolute change in performance would differ (Figure 4.3). For example, if the overall state standard is ninety-five percent,

and program A had an actual performance level of eighty percent in the baseline year, program B had an actual performance level of sixty percent in the baseline year, and both programs were to be held accountable for a ten percent decrease, then program A would be expected to increase by 1.5 percentage points to 81.5 percent, while program B would be expected to increase by 3.5 percentage points to 63.5 percent, after one year. Such a system takes into account the differing base-year levels of performance, and assumes that the closer a program is to standard-level performance, the more difficult it will be to make substantial progress. Over time, both programs would be expected to achieve the same state standard at the same rate, but in differing absolute increments.

Figure 4.3
Performance Targets*



*based on a ten-year improvement cycle

States may develop other systems of performance targets. The most important consideration is to develop a system that recognizes annual improvement efforts, so that low-achieving programs are not penalized for the types of students they serve.

In addition, states should decide what mixture of performance measures will trigger the program improvement process. Should a program that performs poorly on a single measure in relation to the state standard be required to develop a plan or should only those programs that perform poorly on some combination of measures be held accountable? Although performance on some combination of measures may trigger development of the program improvement plan process, programs will probably want to examine performance on each measure separately in order to diagnose their performance problems.

States should decide what mixture of performance measures will trigger the program improvement process.

Finally, states should make performance information readily available. They should provide a range of information to local programs concerning the system of measures and standards, including guidance in developing program improvement plans, and performance reports that indicate how a program performed relative to others in the state. Furthermore, information on performance should probably include the state standards, information on the performance of comparable programs, information on performance in comparison with previous years, and any analysis of the types of programs, schools, or institutions that either exceed, meet, or do not meet standards. In addition, this information should also be made available to students, parents, and industry, as well as other interested and affected parties.

5. Adjustments and Incentives for Special Populations

Section 115 of the Perkins Act specifies that the accountability systems will include "incentives and adjustments that are designed to encourage service to targeted groups or special populations." Neither the law nor the regulations indicate what types of incentives and adjustments are appropriate, and, as with many other aspects of accountability, states enjoy considerable discretion. What "incentives and adjustments" should be made? This chapter discusses a variety of alternatives. It begins by establishing some general principles that should guide the development of incentives and adjustments. Subsequently, it examines several approaches to making programmatic adjustments, as well as modifications to certain types of standards. Finally, it concludes by examining different ways to incorporate incentives into accountability systems.

"Each system . . . will include . . . incentives or adjustments that are . . . designed to encourage service to targeted groups or special populations."

PERKINS ACT
Section 115

Why are incentives and adjustments necessary at all? The primary reason is to ensure fairness in comparisons of relative effectiveness of different programs, secondary schools or postsecondary institutions, or local education agencies. Section 115 specifies that the measures and standards are to be used to conduct annual evaluations of vocational programs. Recipients of federal funds that are not making substantial progress toward meeting the standards established for the state are required to develop local program improvement plans. If after one more year, substantial progress is still not being made, the state is required to intervene. Hence, failing to make substantial progress carries a rather heavy administrative and regulatory burden.

However, because members of special populations are more heavily concentrated in some programs and institutions than in others, differences among programs in meeting standards may not reflect differences in program effectiveness. Instead, some programs face a much more difficult educational challenge than others. If this difference is not recognized through some method of incentives or adjustments, it is possible that programs with higher concentrations of special populations will be subjected unfairly to program improvement requirements. Programs that are subjected to unfair treatment would then have strong incentives to discourage members of special populations from participating. Avoiding this unfortunate outcome and instead encouraging participation are the primary purposes of incentives and adjustments.

General Principles

As a general rule, the same performance measures and standards should apply to *all* students. If measures and standards have validity in the workplace, the failure to meet necessary standards of attainment is likely to lead to failure on the job. Setting lower standards for special populations,

As a general rule, the same performance measures and standards should apply to all students.

therefore, will not help these students. On the contrary, it will only mislead them that they are prepared for jobs that they cannot, in fact, perform. Consequently, lowering standards should be scrupulously avoided. In some instances, lower expectations for attainment by students with special needs are appropriate; however, they are the exception, not the rule. For most special populations, there are better approaches to encouraging participation, completion, and successful job placement.

One of the best approaches is programmatic adjustments that can be made to promote access and successful program completion. Students with special needs place extra demands on teachers. For example, they may require more of a teacher's attention in class, and teachers may find it necessary to modify curriculum, equipment, or assessment instruments. If programmatic adjustments—such as reductions in class size, the addition of a teacher aide, or providing special services—are not made, teachers will be reluctant to admit students with special needs to their programs, especially if teachers are accountable for helping students meet high standards for performance. Effective strategies for programmatic adjustment, at least, make teachers neutral about including students with special needs in their classes. At best, they make teachers enthusiastic and help them create a classroom environment that ensures the success of special populations.

Programmatic adjustments, however, are not always sufficient to ensure that students complete a program in approximately the same amount of time. Consequently, it may be desirable to adjust standards that are defined in terms of rates rather than absolute attainment. For instance, consider two standards for program completion. The first, an attainment standard, specifies that to be considered a program completer, a student must score eighty or higher on a performance-based test of occupational and academic competencies. Assume that this score has been validated in the workplace. Anything lower means a student is not likely to perform effectively on the job and, hence, should not be considered to have completed the program. This example assumes that, by definition, program completion means that a student is ready and able to perform successfully on the job. A second standard, a rate measure, stipulates that programs will be considered to be performing adequately as long as seventy-five percent of those who enter complete the program within two years.

Unnecessary delays in completing a program can be costly for both students and schools. Students who progress slowly through a program—starting, stopping, and restarting—waste time relearning what they have forgotten. They may also be out of sync with the curriculum. They not only take longer to complete but also are likely to require above average time in the program, raising costs per student. Consequently, rate standards can be good policy.

However, some students with special needs simply cannot complete within the time period expected for other students, regardless of whatever programmatic adjustments are made. Presumably they can complete, so modifying the attainment standard would be inappropriate. Yet, they do need more time. Therefore, if teachers are not to be discouraged from including these students in programs, an adjustment in the rate standard

must be made. How such adjustments can be made is explained in detail below.

Finally, there will be instances where some students with special needs cannot meet attainment standards and complete a program, no matter what kinds of programmatic adjustments are made or how much time is allowed. The number of such students is probably much smaller than many people believe, but many schools do serve some of these students. There are, at least, two approaches to handling these students in an accountability system.

First, one could choose to exclude these students from the accountability system. Since they cannot, in fact, reach the validated standards of attainment for the workplace, it makes no sense to hold schools accountable for a result that is impossible to achieve. Exclusion, therefore, would eliminate disincentives to include these students in vocational education programs. It does not create any incentives.

Consequently, a second approach modifies standards for these hardest to serve students who cannot attain the levels of performance expected of everyone else. Lowering standards for certain students is a serious last resort procedure. For the most part, it will be appropriate only for students in special education and a relatively small number of special education students at that. Any lowering of standards for selected special education students should be done in consultation with a special education teacher and should be an official part of the Individualized Education Plan (IEP) that is required for all special education students. Ideally, such changes should be accompanied by an employment strategy that will allow the student to contribute as much as possible in a work setting. Usually such strategies involve careful coordination with an employer to alter a job or occupation to accommodate the special education student.

Any lowering of standards for selected special education students should be done in consultation with a special education teacher and should be an official part of the IEP.

In summary, accountability systems should emphasize high expectations for all students. To the extent that adjustments are necessary to encourage service to members of special populations, programmatic modifications should be made first. Second, if adjustments to standards are necessary, changes in standards that establish rates for assessing program effectiveness should take precedence over changing standards that establish levels of student attainment. Finally, if it can be clearly established that, because of their special circumstances, some students are unable to achieve the desired levels of attainment expected of other students, attainment standards can be adjusted. Lowering standards should be limited to very small numbers of students with the changes incorporated into each student's IEP.

Programmatic Adjustments

Modifying programs to better serve special populations is well understood by vocational educators. Federal policy has for some time encouraged recipients of federal money to use funds to support a variety of programmatic interventions. These have included special counseling, tutoring, curriculum

enhancements, modified equipment and tools, teacher aides, student stipends, child care, and a variety of other services. Such programmatic adjustments are well known, and this chapter will not dwell on them. However, one other adjustment, reductions in class size, does deserve some more discussion.

In many instances, what special needs students require in order to succeed in a vocational program is simply more time with the vocational teacher. However, more time is often the most difficult service for a teacher to provide. Other things being equal, giving more time to some students usually means spending less with others. In extreme cases, the success of the many can be jeopardized by the needs of a few. One strategy for avoiding this dilemma is to reduce class size as the numbers of students with special needs in a class or the severity of their needs increase.

Federal vocational education money can be used to reduce class size for special populations.

At first glance, using federal vocational education money to reduce class size may not seem to be a very good use of those funds. Federal money averages approximately five to ten percent of total expenditures for vocational education. The reduction that could be achieved on the average seems small, even if all the money were used to reduce class size. However, federal money is to be targeted on a relatively small number of eligible recipients with relatively high concentrations of special needs students. For these recipients, the amount of federal money per special needs student can be substantial.

For example, suppose that federal funds are targeted on recipients representing one-fourth of the state's total secondary and postsecondary vocational education enrollment. Assume further that among these actual recipients, special needs students represent one-half of the total student population. If in this example federal funds statewide represented a ten percent increase in total spending for vocational education, they would equal about an eighty percent increase in the expenditures of the actual recipients of federal funds. Or, in other words, with federal funds, the actual recipients would be able to spend eighty percent more per student on special needs students.

Suppose one-half of the federal funds were used to reduce class size in the recipients' vocational programs. Then recipients would be able to spend forty percent more per special needs student. In effect, for purposes of determining class size, a special needs student would carry a weight of 1.4. If the average class size in a vocational program were twenty and if one-half of these students had special needs, class size could be reduced from twenty to sixteen (a 20% reduction) with one-half of the federal allocation. The other half could be used for other types of programmatic interventions.

There are some risks in using federal money to reduce class size. Secondary schools and postsecondary institutions are often reluctant to use "soft" money to support faculty salaries. Changes in federal policy or in the level of federal funding can introduce uncertainties that make it difficult to attract and retain capable staff. Nevertheless, the new targeting provisions of the Perkins legislation will be in effect at least through the 1995–1996 academic year, and some recipients may want to consider the opportunity to reduce class size for special populations as one use of federal funds.

Adjustments to Standards

Ideally, programmatic adjustments would be all that is needed to achieve the desired improvement in program outcomes for members of special populations. In reality, however, these adjustments are often insufficient. Consequently, if programs with different concentrations of special needs students are to be evaluated fairly, some method is needed to recognize that they face different degrees of difficulty in educating students. There are several approaches to ensuring that, for purposes of evaluation, "like" programs are being compared. We will elaborate on three: (1) single indices, (2) composite indices, and (3) weighted outcomes.

If programs with different concentrations of special needs students are to be evaluated fairly, some method is needed to recognise that they face different degrees of difficulty in educating students.

Single Indices

Assume that a state has adopted the following standard: within two years of having begun a secondary or postsecondary program of vocational education, at least seventy-five percent of those entering the program will have successfully completed it.[1] Assume further that there are one-hundred different vocational education programs operating in this state and that, among these, the concentration of students with limited English proficiency ranges from zero to fifty percent. Finally, assume that clear evidence demonstrates that students with limited English proficiency generally need more than two years to complete a program. Under such circumstances, it is unfair to hold all programs to the same program completion standard. Rates may differ not because some programs are more effective than others, but rather, because some have lower concentrations of students with limited English proficiency. Consequently, what standard is appropriate for programs with different concentrations of students with limited English proficiency?

One approach to answering this question is to create an index of limited English proficiency and to rank each program from low to high (or high to low—it does not matter which direction the ranking follows) on this index. Table 5.1 illustrates this approach with some hypothetical data. Along with each program's measure on the limited English proficiency index is its corresponding program completion rate.

[1] Different standards could, of course, be established for secondary and postsecondary programs. For purposes of illustrating this methodology, however, we will assume that the same standard applies to both levels and that all programs are evaluated together.

Table 5.1
Program Completion for Programs by Percentage of Limited English Proficient Students

Shading indicates outlying programs based on the running average.

Rank	Program	LEP Index	Completion Rate	Running Average	Difference
1		0.0	90	87.3	2.7
2		0.5	92	87.4	4.6
3		1.0	91	87.8	3.3
4		1.5	85	87.4	-2.4
5		2.0	89	87.1	1.9
6		2.5	77	87.1	-10.1
7		3.0	88	85.7	2.3
8		3.5	90	85.4	4.6
9		4.0	85	84.8	0.2
10		4.5	84	84.7	-0.7
11		5.0	87	84.2	2.8
12		5.5	75	85.1	-10.1
13		6.0	88	84.7	3.3
14		6.5	85	84.1	0.9
15		7.0	84	84.1	-0.1
16		7.5	83	84.5	-1.5
17		8.0	87	83.3	3.7
18		8.5	84	84.1	-0.1
19		9.0	83	83.6	-0.6
20		9.5	85	83.5	1.5
21		10.0	89	83.4	5.6
22		10.5	73	83.2	-10.2
23		11.0	84	82.5	1.5
24		11.5	83	82.2	0.8
25		12.0	84	81.8	2.2
26		12.5	82	80.6	1.4
27		13.0	81	79.6	1.4
28		13.5	80	80.1	-0.1
29		14.0	80	79.5	0.5
30		14.5	79	78.8	0.2
31		15.0	72	78.2	-6.2
32		15.5	78	77.5	0.5
33		16.0	78	77.0	1.0
34		16.5	77	76.5	0.5
35		17.0	76	76.1	-0.1
36		17.5	77	76.0	1.0
37		18.0	74	75.2	-1.2
38		18.5	76	75.0	1.0
39		19.0	75	75.0	0.0
40		19.5	75	75.3	-0.3
41		20.0	78	75.2	2.8
42		20.5	63	74.9	-11.9
43		21.0	76	74.8	1.2
44		21.5	78	74.6	3.4
45		22.0	80	74.5	5.5
46		22.5	75	74.2	0.8
47		23.0	74	73.5	0.5
48		23.5	73	74.2	-1.2
49		24.0	74	74.0	0.0
50		24.5	73	73.5	-0.5

Table 5.1
Program Completion for Programs by Percentage of Limited English Proficient Students—Continued

Shading indicates outlying programs based on the running average.

Rank	Program	LEP Index	Completion Rate	Running Average	Difference
51		25.0	72	72.8	-0.8
52		25.5	70	72.5	-2.5
53		26.0	71	71.2	-0.2
54		26.5	74	71.6	2.4
55		27.0	73	71.3	1.7
56		27.5	72	71.1	0.9
57		28.0	71	70.8	0.2
58		28.5	60	70.7	-10.7
59		29.0	78	70.5	7.5
60		29.5	70	69.8	0.2
61		30.0	71	69.2	1.8
62		30.5	69	68.7	0.3
63		31.0	69	68.5	0.5
64		31.5	68	68.9	-0.9
65		32.0	67	67.2	-0.2
66		32.5	66	66.6	-0.6
67		33.0	67	65.9	1.1
68		33.5	68	65.3	2.7
69		34.0	65	64.5	0.5
70		34.5	59	63.1	-4.1
71		35.0	64	62.5	1.5
72		35.5	63	62.1	0.9
73		36.0	62	61.9	0.1
74		36.5	61	61.7	-0.7
75		37.0	52	61.5	-9.5
76		37.5	60	61.6	-1.6
77		38.0	62	61.3	0.7
78		38.5	65	60.9	4.1
79		39.0	66	59.8	6.2
80		39.5	62	59.6	2.4
81		40.0	61	60.2	0.8
82		40.5	60	60.0	0.0
83		41.0	59	59.5	-0.5
84		41.5	50	58.1	-8.1
85		42.0	59	57.2	1.8
86		42.5	58	56.6	1.4
87		43.0	58	56.5	1.5
88		43.5	57	56.2	0.8
89		44.0	49	55.8	-6.8
90		44.5	56	55.9	0.1
91		45.0	56	55.5	0.5
92		45.5	59	55.3	3.7
93		46.0	57	54.9	2.1
94		46.5	55	54.9	0.1
95		47.0	51	55.5	-4.5
96		47.5	55	55.5	-0.5
97		48.0	55	55.4	-0.4
98		48.5	54	55.0	-1.0
99		49.0	57	54.7	2.3
100		50.0	56	54.7	1.3

Table 5.1 also computes a running average for each program and displays the difference between the program's completion rate and this running average. The running average is the average program completion rate for the band of programs that include the five programs directly above and below each program. The width of this band is somewhat arbitrary and depends on the total absolute number of programs, as well as the amount of variation on the index within the band. The objective should be to create a band of programs that are more or less like one another on the relevant index. In this case, for any given program, its concentration of limited-English-proficient (LEP) students does not differ by more than plus or minus 2.5 percent from the five programs above and below it. Note that the band must be fewer than eleven programs for the first five and last five programs listed in the table. For example, the band consists of only six programs for the first and last program in the table, because the first program has no programs above it and the last has none below it. The ninth ranked program's running average of 84.8 is the average program completion rate for programs ranked four through fourteen. The program completion rate of program ranked nine is 0.2 percent higher than the running average for the band of the five programs directly above and below it.

The objective should be to define outliers in such a way that there are neither too many nor too few, relative to the total number of programs.

When examining the column listing, the difference between a program's completion rate and its running average allows us to identify "outliers," programs that score significantly higher or lower than the running average. Thus, program ranked six has a program completion rate of seventy-seven percent. This rate exceeds the state standard of seventy-five percent, but it is more than ten percentage points lower than the running average for the five programs on either side of it. Similarly, while none of the programs ranked from sixty to one-hundred have program completion rates that meet the state standard, only three of these programs (seveny-five, eighty-four, and eighty-nine) have completion rates that are more than five percent below the running average.

What should constitute an "outlier?" For purposes of this illustration, we chose programs that had program completion rates that were more than five percent below the running average. This was rather arbitrary, and the choice depends in large part on examining the data directly. The objective should be to define outliers in such a way that there are neither too many nor too few, relative to the total number of programs. As a general rule, if less than ten percent or more than twenty percent of the number of cases are outliers, the definition of an outlier probably should be reconsidered. In this example, the choice of five percent results in thirteen outliers out of a total of one-hundred—nine more than five percent below the running average, and four more than five percent above it.

The table identifies nine programs that had program completion rates that were more than five percent below the running average. Note that two of these had program completion rates that met the state standard. However, relative to other programs like them with respect to concentrations of LEP students, they were significantly less effective. Hence, the state might want to examine these programs more closely to determine whether they should be candidates for program improvement.

This single-index approach has several advantages. First, it provides a relatively simple, easily understood method for adjusting standards. In effect, it says that similar programs will be compared and those programs performing significantly below the average of those most like them will be singled out for program improvement. Second, the approach is easily adapted to different units of analysis. Rather than ranking programs in the fashion illustrated here, the approach would also work for ranking institutions or local education agencies. Additionally, similar tables can be constructed for other types of special populations—for example, special education students and the academically and economically disadvantaged.

The single-index approach says that similar programs will be compared.

The approach, however, also has disadvantages. It is, by definition, one dimensional and ignores the effects on program completion of concentrations of other types of special needs students. For example, the program ranked six in the example above may have scored well below its neighbors because it had an above average concentration of special education students. Composite indices, discussed in more detail below, can help to overcome this problem. Additionally, this approach works only if there are enough observations (e.g., programs, schools, or districts) to make grouping possible and useful. A state with only twenty county school districts would probably not find this method useful for evaluating differences among these districts.

Similarly, the method does not work well if some of the units being compared contain only a few students. For instance, consider a program with only four students, one of whom is LEP. This situation is qualitatively very different from a program of one hundred students, twenty-five of whom are LEP. In both cases, twenty-five percent of the students in the program are limited English proficient, and they will appear as neighbors when ranked. They are not, however, "like" programs.

No strict criterion exists for establishing the minimum size of a unit for using this single index, ranking method. However, as a general rule, we recommend that the method be used very cautiously if it is applied to units containing fewer than twenty-five to thirty students. It should be noted that small numbers always pose a problem for evaluation. It is always dangerous to judge or generalize based on a small number of cases.

Composite Indices

The single-index method does not take into account the multiple effects of differences among programs in the concentrations of more than one type of special population. Other things equal, a program with a high concentration of LEP and special education students faces a greater challenge than a program with only a high concentration of LEP students. A composite index that combines differences on multiple variables into a single measure offers one method for addressing these differences simultaneously.

There are many ways to develop composite indices. Table 5.2 displays a simple method. In addition to the LEP index used in the previous example,

Table 5.2
Program Completion for Programs Ranked by Score on Composite Index

Shading indicates outlying programs based on the running average.

Composite Rank	LEP Rank	Program	LEP Index	Special Ed. Index	Econ. Disad. Index	Acad. Disad. Index	Composite Index	Completion Index	Running Avg.	Difference
1	1		0.0	0.0	0.0	0.0	0.0	90	87.3	2.7
2	2		0.5	0.0	0.0	0.3	0.2	92	87.4	4.6
3	3		1.0	0.0	0.2	0.5	0.4	91	87.8	3.3
4	4		1.5	0.0	0.4	0.8	0.7	85	87.3	-2.3
5	5		2.0	0.2	0.6	1.0	0.9	89	87.1	1.9
6	6		2.5	0.3	0.8	1.3	1.2	77	87.1	-10.1
7	7		3.0	0.5	0.0	1.5	1.2	88	85.7	2.3
8	8		3.5	0.0	1.2	1.8	1.6	90	85.4	4.6
9	10		4.0	0.0	0.3	2.0	1.6	84	84.6	-0.6
10	9		4.5	0.8	0.5	2.3	2.0	85	84.6	0.4
11	11		5.0	1.0	0.7	2.5	2.3	87	84.5	2.5
12	12		5.5	1.0	0.9	2.8	2.5	75	85.1	-10.1
13	13		6.0	2.0	1.1	3.0	3.0	88	84.7	3.3
14	16		6.5	1.6	1.3	3.3	3.2	83	84.1	-1.1
15	14		7.0	1.3	3.0	3.5	3.7	85	84.2	0.8
16	17		7.5	1.5	3.2	3.8	4.0	87	84.5	2.5
17	15		8.0	4.0	1.0	4.0	4.3	84	84.2	-0.2
18	18		8.5	4.2	1.2	4.3	4.5	84	85.0	-1.0
19	19		9.0	4.3	1.4	4.5	4.8	83	83.6	-0.6
20	20		9.5	4.5	1.6	4.8	5.1	85	83.2	1.8
21	21		10.0	4.6	1.8	5.0	5.4	89	83.1	5.9
22	24		10.5	0.0	2.0	5.3	4.4	83	82.6	0.4
23	23		11.0	2.7	3.0	5.5	5.6	84	82.3	1.7
24	22		11.5	5.0	3.2	5.8	6.4	73	82.0	-9.0
25	32		12.0	0.0	0.0	6.0	4.5	78	81.7	-3.7
26	25		12.5	3.0	0.2	6.3	5.5	84	81.2	2.8
27	26		13.0	3.2	0.4	6.5	5.8	82	80.2	1.8
28	28		13.5	3.3	2.0	6.8	6.4	80	79.6	0.4
29	27		14.0	3.5	2.2	7.0	6.7	81	78.5	2.5
30	29		14.5	3.6	2.4	7.3	6.9	80	78.7	1.3
31	30		15.0	3.8	2.6	7.5	7.2	79	78.5	0.5
32	33		15.5	4.2	2.8	7.8	7.6	78	77.9	0.1
33	34		16.0	4.4	3.0	8.0	7.8	77	77.2	-0.2
34	31		16.5	4.5	4.0	8.3	8.3	72	76.7	-4.7
35	39		17.0	0.0	2.0	8.5	6.9	75	76.3	-1.3
36	35		17.5	0.2	4.0	8.8	7.6	76	76.1	-0.1
37	36		18.0	0.3	4.2	9.0	7.9	77	74.3	2.7
38	37		18.5	0.5	4.4	9.3	8.2	74	73.9	0.1
39	40		19.0	5.3	4.6	9.5	9.6	75	72.6	2.4
40	38		19.5	5.5	5.0	9.8	9.9	76	73.4	2.6
41	41		20.0	5.6	5.2	10.0	10.2	78	73.6	4.4
42	70		20.5	0.0	3.0	2.0	6.4	59	73.5	-14.5
43	49		21.0	2.0	2.0	2.3	6.8	74	73.0	1.0
44	42		21.5	2.2	4.0	2.5	7.5	63	73.0	-10.0
45	45		22.0	2.3	3.0	2.8	7.5	80	72.8	7.2
46	44		22.5	5.9	3.2	3.0	8.7	78	72.8	5.2
47	46		23.0	6.1	3.4	3.3	8.9	75	72.4	2.6
48	57		23.5	0.0	2.0	3.5	7.3	71	73.5	-2.5
49	47		24.0	7.0	2.2	3.8	9.2	74	73.0	1.0
50	48		24.5	6.5	2.4	4.0	9.4	73	73.6	-0.6

Table 5.2
Program Completion for Programs Ranked by Score on Composite Index—Continued

Shading indicates outlying programs based on the running average.

Composite Rank	LEP Rank	Program	LEP Index	Special Ed. Index	Econ. Disad. Index	Acad. Disad. Index	Composite Index	Completion Index	Running Avg.	Difference
51	43		25.0	10.0	2.6	4.3	10.5	76	72.8	3.2
52	50		25.5	6.8	2.8	4.5	9.9	73	72.5	0.5
53	51		26.0	7.0	3.0	4.8	10.2	72	71.7	0.3
54	68		26.5	2.0	3.2	10.0	10.4	68	70.3	-2.3
55	52		27.0	2.2	3.4	10.3	10.7	70	69.0	1.0
56	53		27.5	2.3	3.6	10.5	11.0	71	69.0	2.0
57	54		28.0	5.0	5.0	10.8	12.2	74	68.6	5.4
58	65		28.5	3.0	1.0	11.0	10.9	67	68.4	-1.4
59	96		29.0	0.0	0.1	5.0	8.5	55	68.9	-13.9
60	58		29.5	8.0	0.3	5.3	10.8	60	68.6	-8.6
61	55		30.0	7.4	0.5	5.5	10.9	73	68.7	4.3
62	56		30.5	7.6	0.7	5.8	11.1	72	68.5	3.5
63	60		31.0	7.7	2.0	6.0	11.7	70	68.0	2.0
64	59		31.5	7.9	3.0	6.3	12.2	78	67.9	10.1
65	69		32.0	8.0	2.0	17.0	14.8	65	69.2	-4.2
66	61		32.5	8.2	2.2	17.3	15.0	71	69.4	1.6
67	62		33.0	8.3	4.0	17.5	15.7	69	68.8	0.2
68	64		33.5	8.5	3.0	17.8	15.7	68	68.0	0.0
69	66		34.0	9.2	3.2	18.0	16.1	66	67.5	-1.5
70	63		34.5	9.4	3.4	18.3	16.4	69	65.0	4.0
71	73		35.0	4.0	3.6	18.5	15.3	62	64.5	-2.5
72	67		35.5	4.2	3.8	18.8	15.6	67	63.6	3.4
73	72		36.0	4.3	2.0	19.0	15.3	63	63.4	-0.4
74	71		36.5	10.0	5.0	17.5	17.3	64	63.1	0.9
75	95		37.0	3.0	3.0	15.0	14.5	51	62.6	-11.6
76	76		37.5	3.2	2.0	15.3	14.5	60	62.0	-2.0
77	74		38.0	10.4	5.0	15.5	17.2	61	61.1	-0.1
78	79		38.5	10.6	1.0	15.8	16.5	66	60.6	5.4
79	78		39.0	10.7	1.2	16.0	16.7	65	60.4	4.6
80	81		39.5	5.0	6.0	16.3	16.7	61	59.9	1.1
81	80		40.0	5.2	2.0	16.5	15.9	62	59.8	2.2
82	75		40.5	5.3	8.0	16.8	17.6	52	59.6	-7.6
83	77		41.0	15.0	8.2	17.0	20.3	62	59.1	2.9
84	82		41.5	11.0	8.4	17.3	19.5	60	58.5	1.5
85	83		42.0	11.2	8.6	17.5	19.8	59	57.7	1.3
86	84		42.5	11.3	8.8	17.8	20.1	50	57.3	-7.3
87	86		43.0	12.0	5.0	18.0	19.5	58	57.0	1.0
88	94		43.5	4.0	5.2	18.3	17.7	55	57.5	-2.5
89	85		44.0	4.2	5.4	18.5	18.0	59	56.7	2.3
90	88		44.5	13.0	6.0	21.8	21.3	57	56.5	0.5
91	91		45.0	13.2	4.0	22.1	21.1	56	56.1	-0.1
92	92		45.5	13.3	7.0	22.3	22.0	59	55.9	3.1
93	93		46.0	13.5	7.2	22.6	22.3	57	55.7	1.3
94	98		46.5	13.6	8.0	22.8	22.7	54	55.8	-1.8
95	99		47.0	13.8	8.2	23.1	23.0	57	55.7	1.3
96	97		47.5	12.0	11.0	24.0	23.6	55	55.6	-0.6
97	89		48.0	15.0	14.0	24.3	25.3	48	55.6	-7.6
98	100		50.0	15.2	14.2	24.5	26.0	56	55.1	0.9
99	90		50.5	15.3	14.4	24.8	26.2	56	54.9	1.1
100	87		51.0	17.0	10.0	35.0	28.3	58	55.0	3.0

it displays data on indices of special education, economically disadvantaged, and academically disadvantaged students. It calculates a composite index, which in this illustration is simply an average of the four indices. Programs are then reranked on the composite index, with the running average calculated as it was in the previous example. The difference between a program's completion rate and the running average is calculated, and programs with a completion rate that is more than five percent below the running average are identified.

Because there is often a rather high correlation between the concentration of one special population and the concentration of another, one would not expect a composite index to produce dramatically different results from the single-index method. Some programs, however, may shift around. In Table 5.2, this, in fact, occurs. One of the programs (ranked thirty-one on the LEP index and thirty-four on the composite index) that was more than five percent below the running average on the LEP index is no longer below this cutoff when the composite index is used. Three programs (ranked seventy, ninety-five, and ninety-six on the LEP index) now fall below the cutoff on the composite index. In short, one program that the single index identified as substandard is no longer considered as such when the composite index is used. Three programs that were not identified as substandard using the single index are found to need improvement when the composite method is used.

The lower the correlation . . . the more important it becomes to use a composite index to ensure that programs are being assessed equitably.

In most instances, the composite method will yield only small refinements when there is a high correlation among the variables that make up the composite. The lower the correlation, however, the more important it becomes to use a composite index to ensure that programs are being assessed equitably. As with the single-index method, the composite-index method will work less effectively when a small number of units are being compared or a small number of students within units.

Weighted Outcomes

Another approach to compensating for differences among programs in concentrations of special populations is to weight outcomes in the computation of various types of outcome rates. For example, Table 5.3 displays data for fifteen programs. For the period in question, it lists initial program enrollment, the number of special education students completing the program, the number of limited English proficient completers, the number of disadvantaged completers, and the number of other completers. Unweighted completion rates range from sixty-one to ninety-two percent. Assuming that the state standard is seventy-five percent completion, six of the fifteen programs are substandard.

However, assuming that it is more difficult for students with special needs to complete within a given period of time, we now give more weight to a special needs completer than to other completers. In Table 5.3 a special education completer receives a weight of 2.0, an LEP completer a weight of 1.5, and a disadvantaged completer a weight of 1.25. Other completers are not weighted, and each one counts as 1.0.

Table 5.3a
Weighting Program Completion Rates

Program	Initial Enrollment	Special Ed Completers	LEP Completers	Disadvantaged Completers
1	100	2	4	3
2	150	0	1	1
3	125	0	0	0
4	140	10	15	4
5	200	15	2	2
6	175	2	3	1
7	130	0	1	3
8	160	7	8	4
9	165	3	2	1
10	190	20	10	10
11	115	25	5	5
12	195	0	0	0
13	145	10	5	5
14	135	5	5	3
15	170	20	20	10

Table 5.3b
Weighting Program Completion Rates

Other Completers	Unweighted Completers	Unweighted Completion Rate	Weighted Completers	Weighted Completion Rate
70	79	79.00%	83.75	83.75%
100	102	68.00%	102.75	68.50%
115	115	92.00%	115.00	92.00%
60	89	63.57%	107.50	76.79%
120	139	69.50%	155.50	77.75%
130	136	77.71%	139.75	79.86%
100	104	80.00%	105.25	80.96%
100	119	74.38%	131.00	81.88%
125	131	79.39%	135.25	81.97%
90	130	68.42%	157.50	82.89%
40	75	65.22%	103.75	90.22%
160	160	82.05%	160.00	82.05%
100	120	82.76%	133.75	92.24%
70	83	61.48%	91.25	67.59%
100	150	88.24%	182.50	107.35%

Using these weights, we compute a weighted completion rate for each program. Now only two of the programs have weighted completion rates of less than seventy-five percent. Note also that it is possible under this method to have completion rates in excess of one-hundred percent (program 15).

The essential consideration in adopting some type of weighted outcome approach is justification of the weights.

The essential consideration in adopting some type of weighted outcome approach is justification of the weights. Why choose 2.0 for special education students as opposed to 2.5 of 1.25? Clearly, the size of the weight can have a major impact on the weighted completion rate and whether or not a program meets the state standard.

One approach to setting weights is to base them on relative cost. If it costs, on the average, twice as much to help a special education student complete a program, then a weight of 2.0 for special education completers is justified. Alternatively, weights can be based on relative time to completion. For example, if on the average it takes a limited English proficient student one and one-half times as long to complete a program as it does a student with no special needs, a weight of 1.5 is appropriate.

One advantage of the weighted outcome method over the indexing methods discussed above is that it can more easily handle problems with small numbers of units. Using weights to adjust placement rates can be done as easily for ten programs as for one-hundred. Like the indexing methods, however, the weighted outcome approach must be used carefully with units containing small numbers of students. In such units, the addition or subtraction of one student can have a large impact on the weighted outcome rate.

To summarize, indices and weights provide a variety of approaches for adjusting standards to reflect differences among recipients with concentrations of members of special populations. Indices, combined with running averages for bands of "like" units, offer an easy method for ensuring that the results of a particular program or institution are evaluated against others that are more or less like it on relevant measures. This method, however, tends to adopt as acceptable the average effectiveness of a group of similar units. In the event that they are all performing equally poorly, the indexing/running average method for adjusting standards will tend to overlook widespread substandard performance.

Weighting outcomes can be a more accurate method for adjusting standards, provided clear justification for the weights can be established. If so, weighting is much less likely than the indexing/running average method to overlook substandard performance of a group of programs or other comparative units. Ironclad justification for weights, however, may be difficult to achieve.

No method is perfect, and the results of whatever method is chosen must be viewed merely as *indicative* of relative program performance. The quantitative results should be supplemented with more qualitative judgments about which programs (or units) are performing below or above standard. These methods are not the last word on program effectiveness nor are they a

substitute for management and good judgment informed by a variety of evaluative techniques.

Incentives

In addition to adjustments, the Perkins legislation calls for using incentives to encourage services to targeted groups and members of special populations. At a minimum, incentives should discourage providers of vocational education from excluding members of special populations from vocational education. Ideally, they should encourage greater access, completion, and job placement.

At a minimum, incentives should discourage providers of vocational education from excluding members of special populations.

It is common to conceive of incentives as having some financial consequences, but there are incentives that do not have direct fiscal implications. Indeed, given the funding procedures of Perkins, it will be difficult to develop any kinds of fiscal incentives unless states are able to use state funds. Consequently, the discussion here is limited to nonmonetary incentives. These can take at least four different forms: (1) value added measures, (2) standards based on relative improvement, (3) regulatory relief, and (4) awards and recognition. Each of these will be briefly discussed.

Value Added Measures

Recall from the discussion in earlier chapters the importance of including value added measures in the design of accountability systems. An important advantage of including measures of value added or program gains in accountability systems is that such measures can create powerful incentives to serve special populations. If a program is being evaluated on the absolute *gain* in learning that it achieves with students, it is much more likely to achieve the largest gains with students who enter the program initially achieving at very low levels than with students who are already high achievers. In contrast, if it is being evaluated on the number of students attaining an absolute level of achievement, it is most likely to accomplish this aim by admitting students who are already performing at or above the standard.

Relative Improvement

Another way to introduce incentives into accountability systems relies on standards that are stated in terms of relative improvement. For example, suppose that a state establishes a standard for program completion of seventy-five percent or better. Rather than specifying that all programs operating below this standard are poor performers and subject to program improvement plans, the state can set standards for improvement, which if met will be considered sufficient progress.

To illustrate how this method works, assume an overall state standard for completion of seventy-five percent. Assume further that the state will allow

local providers to take five years to reach this standard if they have programs that are not presently meeting it. Programs with completion rates of less than seventy-five percent would divide by five the difference between seventy-five percent and their completion rates. The result would be the annual progress expected from the program. For instance, if a program had a completion rate of fifty-five percent, it would have to show improvement of four percent per year ((75-55)/5=20/5=4).

There are many variants of this method. A state could establish minimum annual increases so that those programs operating closer to the state standard would reach the standard in fewer than five years. Programs operating below the standards could be expected to progress at a rate of at least three percent or one-fifth of the difference, whichever is greater.

Regulatory Relief

Another incentive that can become part of an accountability system is regulatory relief for those making sufficient progress toward state standards. The Perkins legislation already incorporates such a feature in that it relieves districts from developing local program improvement plans if they are making substantial progress. Because the plan can represent a significant administrative burden, local districts have an incentive to sustain sufficient progress. States could significantly strengthen this incentive by not only relieving districts of the federal regulatory requirement but also by relieving them of certain state requirements.

Awards and Recognition

Finally, states may want to incorporate various awards and recognition for high performers—those performing well above state standards, as well as those making above average progress toward meeting state goals. Awards are probably not a very strong incentive, but they should not be overlooked. They are low cost and can help states publicize accountability and program improvement.

Improving access and success for special populations is the primary objective of federal legislation—it is essential that accountability systems be designed to promote this aim.

To summarize, adjustments and incentives for special populations present some of the most important challenges in designing an effective accountability system for secondary and postsecondary vocational education. On the one hand, improving access and success for special populations is the primary objective of federal legislation. It is essential, therefore, that accountability systems be designed to promote this aim. On the other hand, adjusting standards for special populations can do a great disservice if the result is lowered expectations or tacit acceptance of substandard performance. Lowering standards, therefore, must be closely scrutinized and require clear, well-reasoned justification. The methods described in this chapter should help ensure that states maintain high standards for all students while simultaneously encouraging programs to serve and succeed with students who have special needs.

6. Implementation Issues

The Perkins Act calls for implementing systems of performance measures and standards by the Fall of 1992. By this date, states must have developed systems that meet the minimum requirements of Perkins. They will need to begin collecting data on at least two sets of measures, one measuring gains in academic and occupational competencies and another monitoring performance on one of the four other outcomes specified by the legislation. Most states, of course, plan to implement systems that will exceed the minimum requirements of Perkins. Indeed, as we have stressed repeatedly, relying only on the minimum number of measures required by Perkins may introduce undesirable distortions and incentives into the operation of secondary and postsecondary vocational education programs.

For almost all states, performance-based accountability systems will be a new experience—new not only for vocational education but also for education generally. Consequently, the actions that states take to implement these systems in the Fall 1992 must be viewed largely as first steps. In many states, the systems will be incomplete, and in most states, local providers of vocational education will have had no previous experience with performance measures and standards. However, all states will need to continue to monitor and modify their systems as they gain more experience with measures and standards, their role in local evaluations, and their effects on program improvement.

The actions states take to implement these systems in Fall 1992 must be viewed largely as first steps.

Consequently, states will face a number of issues surrounding the implementation of standards and measures. In this chapter, we address three of these issues: (1) opportunities for phasing in accountability systems, (2) kinds of technical assistance that local providers are likely to need, and (3) the need for periodically reviewing and modifying measures and standards over time.

Phasing in Accountability Systems

Although states must be prepared to implement the minimum requirements of Perkins by Fall 1992, those planning expanded systems are free to phase in supplemental activities as they see fit. For example, states may choose to add additional measures at later dates. Alternatively, they may expand their systems to cover greater numbers of students, beginning with those served by federal funds and gradually including more and more participants in vocational education. This section examines some of these opportunities in more detail.

Adding Measures

To meet the minimum requirements of Perkins, states will need at least two sets of measures: (1) one measuring gains in academic and occupational competencies and (2) one of the other four measures required by law (e.g., competency attainment, job or work skill attainment, retention in school, or placement in further education, the military, or employment). As long as states meet these minimum requirements by Fall 1992, they will be free to implement additional measures at later dates.

In many instances, states will not have the necessary assessment tools to adopt certain measures by Fall 1992. Assessing job or work skill attainment may require the development of new types of tests. Obtaining accurate, comprehensive data on placement in the labor market may depend upon making agreements with state agencies responsible for maintaining data on unemployment insurance data, as well as designing appropriate analytic techniques. In many states, multi-state agreements will be required where labor market areas span beyond the borders of a single state. As such efforts are completed, corresponding measures and standards may be added to the accountability system.

States may want to phase in different approaches to measuring gains.

Similarly, states may want to phase in different approaches to measuring gains. Initially, if data are available for only one point in time, it is impossible to measure gains for individual students. Consequently, states will only be able to measure gains programmatically, comparing scores of students in the current year with scores in previous years. As noted in Chapter 3, this approach to measuring gains must include methods for controlling for changes in student characteristics over time so that changes in scores are not inappropriately assumed to be program effects rather than the results of different student demographics. If states subsequently introduce more frequent means of assessment, they may choose to adopt alternative measures of gains. In short, states can adopt programmatic measures of gains for Fall 1992, and plan to phase in individual measures at later dates.

Adding Students

At the time of writing this guide, the U.S. Department of Education had not yet issued final regulations for the Perkins Act. A major unresolved issue in the regulations was to identify the population for whom data on measures and standards would be required. One possibility was that data on measures and standards would be required only for those students who were served by federal funds for vocational education. However, another possibility was that data would be required for the entire vocational education program operated by recipients of federal funds.

Whatever the minimum requirements of the regulations on coverage of accountability systems, most states have indicated that they want their systems to apply to the state's entire program of secondary and postsecondary

vocational education. Many may choose to apply their system statewide in the Fall 1992. Nevertheless, states could choose to phase in ever larger populations of students over time.

For example, in the Fall 1992, a state could begin applying accountability requirements only to those students served by federal funds (if this is, in fact, what the regulations specify). In subsequent years, states could expand the system to include, first, all programs operated by recipients of federal funds and, later, all eligible recipients.

Similarly, states might begin with a rather narrow definition of students to be included in the accountability system. They might choose to limit data collection to those students completing six or more units or credits in vocational education. Because it is impossible to know prospectively how many units or credits students will complete, data on academic and occupational gains would need to be collected for all prospective vocational education students. However, data on attainment or placement could be collected on a smaller subset of students, depending on actual participation over time. As an alternative, states might elect to collect full data only on completers of secondary programs or postsecondary associate degrees. Subsequently, as states gained experience with data collection on these subsets of students, they could expand data collection to include larger groups.

At the postsecondary level, states may want initially to limit accountability to certain classifications of students—for example, full-time students or students who have indicated that their educational objective is a vocational major, degree, or certificate. Then, at a later date, the system could be expanded to include other students participating in the vocational education curriculum. To some extent, of course, a state's ability to adopt such restrictions will depend on how it spends its federal funds.

Adding Programs

At the outset, states may choose to apply measures and standards to vocational education very generally, making little or no attempt to collect and report data separately for particular types of programs. Over time, however, they may want to phase in greater programmatic detail.[1]

Thus, states may elect at a later date to ask local providers to begin reporting data for each major program area—for example, business, marketing, health, occupational home economics, trade and industry, and technical programs. Alternatively, they might ask local providers to report data separately for all programs exceeding a specified number of students—for instance, thirty or more completers in a given year.

[1] The ability to add more programmatic detail will be limited by the number of students participating in programs. Generally, states should be cautious about asking local providers to provide data for groups made up of fewer than twenty-five to thirty students.

States need not implement full-blown, comprehensive systems of measures and standards by Fall of 1992.

In summary, states need not implement full-blown, comprehensive systems of measures and standards by the Fall 1992. Although they must satisfy the minimum requirements of Perkins, they will enjoy considerable discretion in phasing in additions and refinements over time.

Technical Assistance

A major consideration surrounding the implementation of performance measures and standards is technical assistance to local providers of vocational education. Just as performance-based accountability has been new to most states, it will be new to most local systems. States must be prepared to educate local faculty, administrators, and boards about accountability—its purposes, its components, and the expected outcomes. To some extent, the magnitude of this task will depend, in part, on how well states have involved local providers in the development of the accountability systems that are to be implemented in Fall 1992. Nevertheless, almost all states can expect that some education about the basics of accountability will be necessary.

Additionally, local providers are certain to need technical assistance in at least three other areas: (1) performance assessment, (2) technological alternatives, and (3) analysis and use of performance data. Each of these will be briefly elaborated.

Performance Assessment

While most state accountability systems are likely to rely, at least initially, on data that are already being collected at the local level, local providers may need help in extracting data and learning how to use the information in new ways. For example, probably all secondary and postsecondary systems maintain student transcripts. However, local faculty and administrators are not accustomed to using transcripts to assess student performance in any systematic way. On the contrary, teachers and administrators tend to view transcripts simply as administrative records. In a similar vein, local providers use attendance data primarily to obtain state reimbursements; information on attendance has rarely been linked to program evaluation. Even test data, which teachers do use as a means of tracking *student* performance, are rarely used to evaluate *programs*. Hence, it is likely that local providers will need assistance in understanding how they can use data that they have routinely collected for some time to monitor program performance more effectively.

Over the longer term, many states may choose to adopt new techniques for evaluating performance, especially new forms of authentic assessment that will supplement or, in some cases, replace conventional paper and pencil, multiple-choice tests. Portfolios, performance testing, simulated testing, and embedded assessment are some of the innovative ideas that some states and localities are exploring to improve performance assessment. Adopting

these methods will require major staff development efforts to train faculty in using and interpreting the results of these new kinds of assessment.

Technology

For most secondary and postsecondary institutions, implementing performance measures and standards will require automated record systems for collecting, storing, analyzing, and reporting data. Systems will probably function best if individual student records are locally maintained, rather than aggregating student data to some higher level, such as the classroom, program, or institution. Student record data, along with unique student identifiers such as social security numbers, will permit merging student records with other data sets (unemployment insurance data, for example).

Many secondary and postsecondary providers of vocational education do not currently possess the hardware and software needed to support fully automated student record systems.[2] They will need assistance in developing systems that will support performance measures and standards.

Analysis

Most secondary and many postsecondary institutions are unaccustomed to using data to analyze program effectiveness.[3] Traditionally, data have been maintained principally to support institutional archives, to comply with state and federal law, or to obtain state and federal aid. Using data to manage and improve programs, academic or vocational, is relatively foreign to many educators and administrators.

Using data to manage and improve programs, academic or vocational, is relatively foreign to many educators and administrators.

Consequently, many local providers will need assistance in learning how to analyze data to evaluate their vocational programs. Analytic issues of particular importance include the following:

- Using transcripts to assess student course-taking patterns, especially progress in pursuing and completing a coherent sequence of academic and vocational studies;
- Using longitudinal and cross-sectional methods to measure student gains in academic and occupational skills;
- Controlling for differences in student demographics when analyzing changes in student outcomes attributable to programmatic changes;
- Conducting methodologically sound student follow-up; and

[2] Postsecondary institutions are more likely to have automated systems than secondary institutions; however, it is likely that some postsecondary institutions will need technical assistance with system development.

[3] Many postsecondary institutions do support institutional research units, but these are far from universal. These research units have also been the victim of many budget cuts and may not be as numerous as they were five to ten years ago.

- Presenting analytic results clearly and simply so that they are easily understood by students, parents, teachers, administrators, and policymakers.

In summary, states can expect substantial demands for technical assistance in the first year or two of implementing performance measures and standards. Unfortunately, these demands will occur at a time when many states are experiencing severe cuts in state staff and other funding for providing help to locals. Consequently, states may find it expedient to work together to develop joint technical assistance efforts.

Monitoring and Modification

States would be wise now to provide for annual review of their accountability systems.

Developing systems of performance measures and standards is a complicated and, at times, uncertain process. The systems that states implement in Fall 1992 will be far from perfect and will require ongoing monitoring and modifying as experience is gained at both the state and local levels. States would be well advised now to provide for annual review of their accountability systems. These reviews could take several forms, including an ongoing role for the Committee of Practitioners, which is the group charged with advising states on the development of accountability systems.

Annual Review

States should annually review their accountability systems to assess how well they are working and what kinds of changes are needed. These annual reviews should address several issues.

First, the system should be evaluated against the six basic requirements set forth in the first chapter of this guide:

1. The desired measures and standards should be clearly and precisely defined.
2. Outcomes should be easily and accurately measured, while minimizing data burden and cost.
3. A manageable number of outcome measures should be included in the accountability system.
4. Standards must have validity and be justified in terms of success in either the work place or further education and training; they must also be fair, avoiding bias by race, gender, or special need.
5. Data for each measure should be collected at appropriate intervals of time.
6. The information generated by the accountability system should be routinely accessible to students, teachers, administrators, parents, employers, board members, and others interested in educational policy and performance.

The first question that states should consider during the annual review is the following: How well has the system satisfied these requirements during the previous year of operation?

Second, are there any signs that the system is producing undesirable effects on program operations? Of particular concern is evidence of creaming or other kinds of activities that might limit the access and success of members of special populations. The effects of adjustments and incentives should also be examined annually.

Third, do the performance measures need modification and are changes in the number of measures desirable? The review should consider the extent to which the system may be encouraging undue emphasis on some student outcomes, while excluding others. Conversely, it is possible that the system needs more focus and that the number of measures should be reduced. Additionally, the review should examine whether definitions of performance measures need clarification or change.

Fourth, are the performance standards appropriate? Are some standards too high or too low and on what basis can this be determined? The review should also assess the validity of the established standards, with special attention to their relevance to the workplace or further education.

Fifth, how are local providers performing? Are there any clear patterns with respect to below average or above average performance? At this time, criteria for what constitutes sufficient progress toward meeting state standards should be reviewed. The state should examine what types of providers are not making sufficient progress and why.

Sixth, what kinds of problems are local providers experiencing in operating and using the accountability system? The state should continue to assess needs for technical assistance at the local level.

Finally, the state should continue to review the role of the performance measures and standards for vocational education in the larger context of general accountability for all of secondary and postsecondary education. Generally, states will be better served by integrated systems of assessment and evaluation that help to examine the relationship and contribution of vocational education to the larger education system.

States will be better served by integrated systems of assessment and evaluation.

Ongoing Role for Committee of Practitioners

States can conduct this annual review in a number of ways. State staff could conduct the review, as could independent consultants. However, one approach that states should consider is continuing to rely on the Committee of Practitioners to fulfill this role.

The federal legislation requires that states establish a Committee of Practitioners to advise the state in the development of accountability systems. While the Congress seems to have envisioned the Committee as the functioning body during the period prior to implementation in Fall 1992, states may continue to rely on this group for ongoing development of measures and standards.

States could convene the Committee in late spring or early summer to review the most recent experience with the accountability system. Work by state staff and independent consultants could supplement the Committee's review. As was the case for the Committee's role under the federal legislation, the Committee's responsibilities would continue to be advisory. It would make recommendations to state staff and the State Board of Education, and state staff and the State Board would be free to act on these recommendations as they see fit.

Assuming the Committee of Practitioners is maintained by the state, states should ensure that these Committees include sufficient representation by employers and labor. Although it was probably an oversight, the Perkins Act failed to include employers and labor in the list of groups to be represented on the Committee. This should be corrected, especially as states seek to further validate the performance standards adopted as part of the accountability system.

Implementation of measures and standards in Fall of 1992 is only the beginning in an ongoing process of system development and refinement.

In summary, implementation of measures and standards in Fall 1992 is only the beginning of an ongoing process of system development and refinement. Of necessity, many states will begin by implementing only the minimum requirements of federal legislation. In some cases, meeting even the minimum requirements of Perkins may prove problematic. Effectively measuring academic gains at the postsecondary level, for example, would require major developments in postsecondary testing and assessment in many states. By providing a process for ongoing review and modification, states can ensure that accountability will become an increasingly useful tool in improving secondary and postsecondary vocational education.

7. Student Assessment

In order to provide information on the attainment of both academic and occupational skills, states will need to employ one or more types of assessment instruments in their performance measures and standards systems. With the thousands of tests available and the many conflicting assessment philosophies, choosing assessment instruments can be a daunting task. Today conceptions of appropriate methods for assessing student achievement and mastery of competencies are shifting. Because unintended negative effects of standardized multiple-choice tests on student learning have been recognized, these tests are now being replaced or supplemented with more authentic assessments—ones that are expected to affect student learning in positive ways. Most states will initially want to use some form of existing standardized tests, supplemented with a system of authentic assessment to be phased-in over time.

Most states will initially use some form of existing standardized test, phasing in authentic assessment over time.

This chapter provides an introduction to assessment issues in the context of performance measures and standards. Given the complexities of assessment, it is not possible to explore all the issues or technical aspects of assessment here; to do so would require an additional guide. This chapter begins with a review of the terminology and definitions used to describe different types of tests. Next, it presents an overview of the issues and procedures involved in selecting an assessment instrument. The following section describes how assessment is used in the system of performance measures and standards. Finally, this chapter provides a summary of ongoing activities in a sample of states that are currently implementing authentic assessment systems. References are supplied at the end of this chapter for those who want more detailed information about these issues.[1]

Student Assessment Instruments

Assessment instruments differ with respect to many important characteristics such as purpose, methodology, and the inferences that can be drawn from scores. The following section provides a brief discussion of different assessment instruments and the varying characteristics of the instruments.

[1] The references to tests and test development organizations in this chapter are not endorsements by the National Center for Research in Vocational Education. This chapter is designed to provide an overall description of assessment and how it applies to performance measures and standards, not to advocate particular tests or test development organizations.

Norm-Referenced Versus Criterion-Referenced

One way to classify assessment instruments is by their method of referencing results, in other words, by the way in which an individual's performance on an assessment instrument is reported. For example, it is important to ask: is the performance referenced to how other students perform on the instrument, or to a specific set of criteria? Two types of tests are based on the method of referencing test results: *norm-referenced* and *criterion-referenced*. These are described in detail below.

Norm-referenced tests show how an examinee performed in relation to others.

Norm-referenced tests show how an examinee performed in relation to others. Typically these tests feature a fixed set of items along with specific instructions for administration and scoring. Norm-referenced tests are first administered to a representative sample of students (the norm group) to establish norms, which reflect the test score distribution of students in the original sample and provide a frame of reference against which the performance of other students (or groups) can be assessed. From these test results, one can see whether a student performed above or below the norm and by how much. Examples of norm-referenced tests include the *Iowa Test of Basic Skills*, the *Stanford Achievement Test*, and the *Metropolitan Achievement Test*.

Criterion-referenced tests compare results against a specific body of knowledge and skills.

Criterion-referenced tests reference results in terms of specific knowledge and skills; they are constructed to support generalizations about an individual's performance relative to a specified set of tasks. From these test results, one can see how well a student performed relative to the specific standards or criteria of performance. Examples of criterion-referenced tests include the *National Assessment of Educational Progress*, the *Stanford Diagnostic Mathematics Test*, and the *Advanced Placement Exam*.

In the midst of the controversy over the most appropriate methods for assessing student achievement and mastery of competencies, it is difficult to favor norm-referenced tests over criterion-referenced tests for several reasons. With a norm-referenced test, there is no absolute standard against which to measure individual performance, and some argue that examining student performance relative to a norm group creates winners and losers. The argument follows that results should concentrate on what students learn in substantive, criterion terms, and therefore, students' test results should be measured against an expected standard. However, if the objective of the assessment is to see how one student performs compared with others in the nation (or state), a norm-referenced test may be appropriate.

Direct Versus Indirect Assessment

A good way to understand the differences in assessment instruments is with respect to how they measure a skill—directly or indirectly. The distinction between direct and indirect can be thought of as a continuum (Figure 7.1) with norm-referenced, multiple choice exams on one end of the spectrum and performance assessment on the opposite end with numerous assessment instruments positioned at varying points between the two ends of the spectrum.

Figure 7.1

A direct assessment directly measures a skill or task while the student is actually performing or demonstrating it. An indirect assessment instrument is designed to make inferences about what the student can or cannot do without the student actually performing the task. For example, a direct assessment of auto mechanics skills may require the student to actually change the oil of a car. This assessment would directly determine whether the student could or could not perform the task of changing the oil. In contrast, an indirect assessment instrument may involve asking the student multiple-choice questions about the knowledge and procedures involved in changing the oil of a car. From the results of this test, one can then infer whether the student could or could not change the oil of a car if faced with this task.

Assessment tasks differ with respect to how they measure a skill—directly or indirectly.

Types of Student Assessment Instruments

Standardized tests, generally thought of as indirect instruments, are administered and scored under uniform testing conditions to permit comparison among all students (these may either be norm-referenced or criterion-referenced). Although most people associate standardized tests with the multiple-choice format (indirect), standardization is a characteristic that in theory can apply to any test format—from written essays to oral examinations to the production of a portfolio (direct). The goal of standardization is to make test scores comparable and to ensure that test-takers have equal chances to demonstrate what they know.

Competency-based tests are criterion-referenced tests that relate to specific curriculum. A competency-based assessment usually features specific tasks, predetermined standards of performance, a range of difficulty, and are administered frequently for feedback and advancement. This kind of assessment can be standardized in the form of a checklist where the administrator checks off the competencies attained, or as an administered paper-and-pencil test (indirect), or in actual performance-based tasks (direct).

Performance assessment requires students to demonstrate competencies directly, by producing specific information on student mastery of skills, rather than information that indirectly indicates student knowledge or skills. Students must demonstrate a process or produce a product in situations that resemble classroom instruction, with the goal of measuring knowledge-in-use and the students' ability to apply skills that closely approximate those used in real life. For example, paper-and-pencil tests that

Performance assessment requires students to demonstrate competencies.

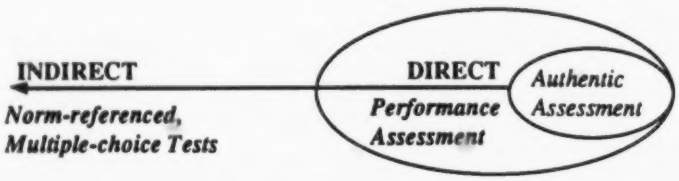
require students to produce the solution to a question rather than to select from an array of possible answers are often considered to be performance assessment. In performance assessment, tasks may be broken down into many small steps, and then students are graded on their ability to perform each small step. On the other hand, it may also involve complex tasks that can be approached in a variety of ways and that are typically scored holistically. Examples of performance-based assessment include the *NOCTI cabinet-making test, essays, oral discourse, and experiments.*

Authentic assessment goes one step beyond performance assessment. It usually requires higher order thinking and problem-solving skills, can involve a wide range of activities that extend over time, and is closely integrated with instruction. In vocational education, authentic assessment integrates academic and vocational skills and can include a self-assessment or a reflective component.

Authentic assessment replicates the challenges and standards of performance that writers, business people, scientists, community leaders, or designers typically face in their work.

Authentic assessment replicates the challenges and standards of performance that writers, business people, scientists, community leaders, or designers typically face in their work. Such activities include *writing essays and reports, conducting individual and group research, designing proposals and mock-ups, assembling portfolios, and many other kinds of performance tasks.*

Figure 7.2



To demonstrate how assessment classifications overlap, consider performance assessment. If a performance-based instrument is in the form of a checklist, in which the instructor checks off items that the student performs successfully, the instrument may be considered a competency-based instrument. In other words, the instructor uses the instrument so that the student can perform specific competencies or individual tasks. If a performance instrument simulates work or a work experience, it may be an authentic assessment tool. The first type of performance-based instrument assesses specific tasks within a job, while the second assesses the whole job or an entire situation.

Evaluating the Assessment Instruments

Assessment instruments are evaluated according to a variety of different criteria, the most important being *validity*, *reliability*, and *cost*.

Validity

Validity refers to whether or not the instrument measures what it is supposed to measure, as well as the appropriateness, meaningfulness, and usefulness of the specific inferences made from test scores. Validity is evaluated by the degree to which evidence supports the inferences that are made from the scores. Thus, the inferences made in specific uses of a test are validated, rather than the test itself. Validity is judged by many types of evidence, including the consequences of translating test-based inferences into decisions or policies that can affect individuals (For more information, see the Recommended Reading list, p. 126).

Reliability

Reliability refers to the degree to which test scores are free of measurement errors and are consistent in their results. For example, even if a reliable test assesses individuals at different times, uses a different assessor, or contains a different sample of equivalent items or tasks, similar results should be obtained. When there are differences between scores from one occasion to another, this may be attributable to what is commonly called *measurement error*, which reduces the reliability (and therefore the generalizability) of the individual's score from a single measurement. The lower the degree of measurement error, the greater the confidence that the test scores accurately represent what is being measured (For more information, see the Recommended Reading list, p. 126).

Cost

Standardized, multiple-choice tests are relatively inexpensive to purchase and implement, yet developing and implementing an authentic assessment system is very costly and time-consuming. In theory, this cost is more than compensated for by the understanding that money invested in an authentic assessment system provides much more than assessment. When handled properly, curriculum and instruction benefit greatly when authentic assessment is implemented because teacher time invested in assessment is no longer taken from instruction. In other words, assessment and instruction become one. A further benefit is that the training required to prepare teachers to participate in the assessment system also turns out to be effective professional development. The California Assessment Program writing assessment, which is widely credited with lifting the state's writing instruction to new levels of excellence, is a good example of such a benefit.

When handled properly, curriculum and instruction benefit greatly when authentic assessment is implemented, because assessment and instruction become one.

The Questions to Ask

In order to determine which particular assessment instrument to use, states should ask themselves specifically what it is they are trying to measure. Answering this question should lead to a particular instrument (or at least narrow down the choices). For example, the Perkins Act calls for a measure of learning and competency gains, including student achievement of basic or more advanced academic skills. If a particular state wanted to choose an instrument that assessed a student's achievement in an advanced academic skill such as writing, the state must first decide specifically what it is trying to measure. If it decides to assess the student's ability to think critically and creatively about a topic and to use grammar rules correctly in essay writing, a standardized, multiple-choice test may not be the first choice. This type of test may only assess knowledge of grammatical rules, not whether the student can actually apply these rules. A multiple-choice test would not assess whether the student could construct an essay (or at what level of performance). Instead, in this instance, the state could choose a combination competency/performance-based test in which the students would be asked to construct an essay on a given topic. The student would then be scored in relation to the desired competencies to determine their individual level of performance. This may require states to begin with a multiple-choice test and phase in a competency/performance-based test over time.

Use of Student Assessments in the System of Performance Measures and Standards

How can assessment instruments be most effectively used in the context of a system of performance measures and standards? For Perkins II, state administrators may choose to use assessment instruments to measure learning outcomes such as basic, academic, and occupational skills (including job-specific, industry-specific, and generic). In most states, at least in secondary education, a standardized, multiple-choice academic test is already in use. Given the previous discussion, these statewide tests could be used to measure basic skills. However, they may not be appropriate for assessing advanced academic skills. Separate instruments, of course, will be needed to measure occupational competencies.

In a work setting, individuals need to apply their knowledge; performance-based assessments simulate that kind of situation.

In vocational education, performance and authentic assessments are particularly appropriate since one goal of vocational education is to enable students to perform their jobs and to use broad, advanced skills to adapt to changes. In a work setting, individuals need to apply their knowledge directly to accomplish goals, solve problems, and respond to new situations. Performance-based assessments simulate this kind of situation much more closely than do standardized tests.

However, it will take some time for states to move from standardized to more performance-based assessments. As Ruth Mitchell reports, the system of education must move from "testing to assessment" (Mitchell, 1992). Whereas a "test" is a single-occasion, unidimensional, timed exercise, usually in multiple-choice or short answer form, an "assessment" is an activity that can take many forms, can extend over time, and aims to capture the quality of students' work or an educational program. Thus, states should dedicate substantial resources and long-term planning to the area of assessment. At the present time authentic assessment, although well-grounded in research to establish its philosophical base, is still in its early stages of development. Despite the fact that many people are developing and piloting authentic assessment nationwide, much work remains to be done with respect to establishing large-scale authentic assessment systems. In fact, statewide assessment tools that truly evaluate the integration of academic and vocational skills are nonexistent at this time.

Much work remains to be done with respect to establishing large-scale authentic assessment systems.

Within their performance measures and standards system, we recommend that states use a multi-indicator system that builds on already implemented standardized tests, while phasing-in performance and authentic assessment systems over time. Reliance on a single measure, taken at a single point in time, cannot represent the complexity of what one needs to know about vocational education students' academic and occupational skills, or about how school and institution programs affect the development of these skills. Similar to our phase-in proposal of performance measures and standards, this multi-indicator system will require phasing-in its components over time.

Beginning a Multi-Indicator Assessment System

1. Determine all the learning outcomes that the state would ideally like to measure in the long-term.
2. Identify the learning outcomes currently being measured.
3. Determine assessment instruments to measure learning outcomes currently not measured in order to supplement ones currently measured.
4. Phase-in according to priorities and resources.

States should start with a vision of what they want to assess in the long run (e.g., in the next five to seven years) and plan components to phase-in each year. (See Chapters on Performance Measures and Standards for a more graphic representation of how this might look.) Consider the following example. A state may want to develop an assessment system that measures specific and generic occupational skills as well as basic and advanced academic skills. This same state has a norm-referenced, standardized test in

States should create a multi-indicator system that builds on already implemented standardized tests, while phasing-in performance and authentic assessment.

place and would like to move toward a multi-indicator system and may also decide to institute a writing assessment this year. This writing assessment might be a timed, situation-specific essay that students would write and teams of teachers would score. Over the next year, the state might implement a statewide employability skills portfolio requirement (collection of student work) that might include, among other items, a resume, a report on an internship, and a letter of recommendation. Finally, they might want to include a work sample in their assessment system.

States should begin looking at what assessments they currently have in place and anticipate what they would like to achieve in the future.

In developing this multi-indicator assessment system, states should follow these rules: (1) develop a long-term plan or a vision specifying the goals for assessment, and (2) *go slowly* by phasing-in assessment components over time. States should begin by looking at what components they currently have in place and then anticipate what they would like to achieve in the future.

Some possible assessment components that states might use to supplement existing standardized tests in their multi-indicator system are listed below.

- Writing Assessment
- Portfolio
- Exit Interview Exam
- Written Scenario Piece
- Work Sample
- Senior Project

State Initiatives

Some states have started to invest considerable time and resources in developing an authentic assessment system. In fact, most statewide initiatives are aimed at reforming academic assessment. These efforts should serve as valuable models and resources for other states interested in reforming or adding to their assessment systems. The next section provides an overview of a few of these state initiatives.

California

California is now in the process of completely revamping its statewide assessment system. The new California Assessment Program (CAP) will include portfolios, projects, writing samples, and a variety of tasks embedded in the curriculum. This new system will be phased-in over a five-year period, while at the same time, multiple-choice testing will be gradually de-emphasized. In addition, the Far West Laboratory is developing a system of authentic student certification assessments for vocational education for California which include portfolios, projects, presentations, and written problem-solving scenarios.

Connecticut

Connecticut is devising individual and group-administered performance exercises in mathematics and science as part of a statewide mastery testing program. The state plans to expand this innovative work to other content areas.

States involved in assessment reform can serve as models and resources for other states.

Michigan

Michigan has developed performance indicators to be used in statewide assessments in art, music, science, mathematics, social studies, and career development. In addition, the state has pilot-tested an employability skills portfolio. This assessment will eventually be administered to all students in grades eight through twelve to document academic, personal management, and teamwork skills and to assess work readiness.

Vermont

Vermont is implementing a portfolio assessment process in writing, science, mathematics, and other areas to be administered on an individual basis to students. These portfolios contain standard assessment exercises for all students, along with components that are selected by individual students. All portfolios are evaluated by teachers throughout the state to ensure that comparable standards are used to evaluate individual students.

Vocational Education Assessment Contacts

A number of vendors and organizations are exploring ways of assessing occupational competencies and academic skills that, like the state initiatives, should serve as valuable models for states.

Advanced Systems in Measurement and Evaluation
171 Watson Road
P.O. Box 1217
Dover, NH 03820
(800) 431-8901

Advanced Systems has a test development group that develops test instruments, scoring criteria for nonmultiple-choice responses, and interpretive reports of results. Maine Educational Assessment, the Massachusetts Basic Skills Testing Program, and the Massachusetts Educational Assessment Program have been their three primary contracts.

American College Testing (ACT)
Work Keys
P.O. Box 168
Iowa City, IA 52243
(319) 337-1000

Work Keys, a new initiative underway at ACT, is a national system for teaching and competency-based assessment of employability skills. Learners entering the Work Keys system will be assessed to identify the performance level at which they demonstrate competencies for each of the Work Keys skills.

Council of Chief State School Officers
One Massachusetts Avenue, NW, Suite 700
Washington, DC 20001-1431
(202) 408-5505

In conjunction with their three-year priority, "Connecting School and Employment," the Council of Chief State School Officers has sponsored a Workplace Readiness Task Force to encourage state departments of education and the business community to work together in developing joint assessment instruments.

States interested in assessment reform can contact a variety of vendors and organizations in exploring ways of assessing occupational and academic skills.

Educational Testing Service
Rosedale Road
Princeton, NJ 08541
(609) 921-9000

The Educational Testing Service (ETS) has a separate division, the Center for Occupational and Professional Assessment (COPA),that focuses on job-related assessment programs and services. Most COPA tests are multiple-choice, but the ETS also uses other assessment techniques such as essay tests, performance tests, and simulations.

Far West Laboratory
730 Harrison Street
San Francisco, CA 94107-1242
(415) 565-3061

Contracted by the California State Department of Education, Far West Laboratory is developing the California Career-Vocational Education Student Certification Project. This project develops and field-tests an authentic student certification system for vocational education programs offered in California high schools, regional occupational programs/centers, and adult education programs. Student certification is intended to demonstrate that students have mastered vocational and academic competencies consistent with vocational area Model Curriculum Standards, thereby increasing their access to employment and postsecondary opportunities.

Two general categories of assessment are used for all certifications: cumulative and administered. The cumulative assessments include a supervised practical experience, a certification project, and a comprehensive portfolio. The administered assessments include performance tasks, a written scenario, and a multiple-choice test. These various methods provide students with multiple opportunities for success, emphasizing direct over indirect indicators of competence.

National Occupational Competency Testing Institute (NOCTI)
409 Bishop Hall
Ferris State University
901 South State
Big Rapids, MI 49307-2295
(616) 796-4695
(800) 334-6283

The National Occupational Competency Testing Institute provides occupational competency examinations and related services to both the United States and world communities, as well as NOCTI teacher (TOCT), student (SOVAT), and industrial (IOCT) testing products.

Vocational–Technical Education Consortium of States (V-TECS)
1866 Southern Lane
Decatur, GA 30033-4097
(800) 248-7701

Since its founding in 1973, the primary activity of the Vocational–Technical Education Consortium of States (V-TECS) has been to define and analyze occupations for the purpose of creating competency-based curriculum and programs. V-TECS has been developing competency-based assessment instruments based on performance objectives, and sponsors an annual conference, the National Conference on Competency-based Testing and Performance Standards for Vocational–Technical Education. During the last two years, development of the V-TECS Data System, a software package designed for storing and retrieving V-TECS material, has been one of their major efforts.

It should also be mentioned that vocational student organizations are a valuable resource for developing performance assessments. Future Homemaker Association (FHA), Vocational Industrial Club of America (VICA), Future Business Leaders of America (FBLA), Health Occupations Student Association (HOSA), Distributive Education Club of America (DECA), Future Farmers of America (FFA), Technology Student Association (TSA), Young Farmers Association (YFA), Postsecondary Agriculture Student Association (PASA), and others organize national competitions for students that are in essence performance assessments.

Student organizations can serve as resources on performance-based assessment.

Recommended Reading

Anastasi, A. (1988). *Psychological Testing*. New York, NY: MacMillan.

Berlak, H., et al. (1992) *Toward a new science of educational testing and assessment*. Albany, NY: State University of New York Press.

Committee to Develop Standards for Educational and Psychological Testing. (1985). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association.

Congress of the United States Office of Technology Assessment. (1992, February). *Testing in American schools: Asking the right questions* (OTA-SET-520). Washington, DC: U.S. Government Printing Office.

Conoley, J.C., & Kramer, J.J. (Eds.). *Tenth mental measurements yearbook*. Lincoln, NE: University of Nebraska Press.

Cronbach, L.J. (1988). Five perspectives on validity argument. In H. Wainer & H.I. Braun (Eds.), *Test Validity*. Hillsdale, NJ: Erlbaum.

Redirecting assessment. (1989, April). *Educational Leadership, 46* (7).

Feldt, L.S., & Brennan, R.L. Reliability. In R.L. Linn (Ed.)., *Educational Measurement* (pp. 105–146). New York: MacMillan.

Frederiksen, J.R., & Collins, A. (1989, December). A systems approach to educational testing. *Educational Researcher, 18*(9), 27–32.

Gardner, H. (1991). Assessment in context: The alternative to standardized testing. In B. Gifford & M.C. O'Connor (Eds.), *Cognitive Approaches to Assessment*. Boston, MA: Kluwer Academic.

Johnston, P. (1987, April). Teachers as Evaluation Experts. *The Reading Teacher, 40*(8), 744–748.

Linn, R., Baker, E., & Dunbar, S.B. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher, 20*(8), 15–21.

McLaughlin, M.W. (1991). Test-based accountability as a reform strategy. *Phi Delta Kappan, 73*, 248–251.

Messick, S. Validity. (1989). In R.L. Linn (Ed.), *Educational Measurementt* (pp. 13–104). New York, NY: MacMillan.

Mitchell, R. (1992). *Testing for learning: How new approaches to evaluations can improve american schools*. New York, NY: Macmillan Free Press.

Petrone, V. (Ed.). (1991). *Expanding student assessment*. Alexandria, VA: Association for Supervision and Curriculum Development.

Stiggins, R.J. (1987). Design and development of performance assessments. *Educational Measurement: Issues and Practice, 6*(3), 33–42.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan 70*, 703–713.

Wolf, D., Bixby, J., Glenn, J., & Gardner, H. (1990). To use their minds well: Investigating new forms of student assessment. *Review of Research in Education 17*, 31–74.

Glossary of Terms

All aspects of an industry includes—with respect to a particular industry that a student is preparing to enter—planning, management, finance, technical and production skills, underlying principles of technology, labor and community issues, health and safety, and environmental issues related to the industry. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Area vocational education school means (1) A specialized high school used exclusively or principally for the provision of vocational education to individuals who are available for study in preparation for entering the labor market; (2) The department of a high school exclusively or principally used for providing vocational education in not less than five different occupational fields to individuals who are available for study in preparation for entering the labor market; (3) A technical institute or vocational school used exclusively or principally for the provision of vocational education to individuals who have completed or left high school and who are available for study in preparation for entering the labor market; or (4) The department or division of a junior college, community college, or university that operates under the policies of the state board and provides vocational education in not less than five different occupational fields leading to immediate employment but not necessarily leading to a baccalaureate degree, if, in the case of a school, department, or division described in paragraph (3) or in this paragraph, it admits as regular students both individuals who have completed high school and individuals who have left high school. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Coherent sequence of courses means a series of courses in which vocational and academic education are integrated, and which directly relates to, and leads to, both academic and occupational competencies. The term includes competency-based education, academic education, and adult training or retraining that meet these requirements. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Cohort means a group of individuals who have a statistical factor in common.

Competency attainment is referred to in the guide as specific occupational competency attainment. [See *specific occupational competency attainment*.]

Eligible institution means an institution of higher education, a local educational agency serving adults, or an area vocational education school serving adults. [See *eligible recipient*.] (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 403.117(a).)

Eligible recipient means, except as otherwise provided, a local educational agency, an area vocational education school, an intermediate educational agency, a postsecondary educational institution, a state corrections educational agency, or other eligible institution. [See *eligible institution*.] (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

General occupational skills means strong experience in, and understanding of, all aspects of an industry. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Intermediate educational agency means a combination of school districts or counties that are recognized in a state as an administrative agency for that state's vocational or technical education schools or for vocational programs within its public elementary or secondary schools. This term includes any other public institution or agency having administrative control and direction over a public elementary or secondary school. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Job or work skill attainment or enhancement is referred to in the guide as general occupational skills attainment or enhancement. [See *general occupational skills*.]

Local educational agency means a board of education or other legally constituted local school authority having administrative control and direction of public elementary or secondary schools in a city, county, township, school district, or political subdivision in a state, or any other public educational institution or agency having administrative control and direction of a vocational education program. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Measure means a description of an outcome. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Outcome means a measurable aspect of student performance.

Postsecondary educational institution means an institution legally authorized to provide postsecondary education within a state, a Bureau of Indian Affairs-controlled postsecondary institution, or any postsecondary educational institution operated by or on behalf of any Indian tribe. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Sequential course of study means an integrated series of courses that are directly related to the educational and occupational skills preparation of individuals for jobs, or preparation for postsecondary education. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Special populations include individuals with handicaps, educationally and economically disadvantaged individuals (including foster children), individuals of limited English proficiency, individuals who participate in programs designed to eliminate sex bias, and individuals in correctional institutions. (Carl D. Perkins Vocational and Applied Technology Act of 1990, Title V, Part C, § 521(31).)

Specific occupational competency attainment refers in the guide to Title II occupational programs.

Standard means the level or rate of an outcome. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Technology education means an applied discipline designed to promote technological literacy that provides knowledge and understanding of the impacts of technology including its organizations, techniques, tools, and skills to solve practical problems and extend human capabilities in areas such as construction, manufacturing, communication, transportation, power, and energy. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

Tech-Prep education program means a combined secondary and postsecondary program that (1) leads to an associate degree or 2-year certificate; (2) provides technical preparation in at least one field of engineering technology, applied science, mechanical, industrial, or practical art or trade, or agriculture, health, or business; (3) builds student competence in mathematics, science, and communications (including through applied academics) through a sequential course of study; and (4) leads to placement in employment. (Carl D. Perkins Vocational and Applied Technology Act of 1990, Title III, Part E, § 347(3).)

Title III programs include special programs such as Vocational Education Support Programs by Community-Based Organizations, Consumer and Homemaking Education, Career Guidance and Counseling Programs, Business-Labor-Education Partnership for Training, Tech-Prep Education, State Grants for Facilities and Equipment, Community Education Employment Centers and Vocational Education Lighthouse Schools, and Tribally Controlled Postsecondary Vocational Institutions.

Vocational education means organized educational programs offering a sequence of courses, or instruction in a sequence, or aggregation of occupational competencies that are directly related to the preparation of individuals for paid or unpaid employment in current or emerging occupations requiring other than a baccalaureate or advanced degree. These programs must include competency-based applied learning that contributes to an individual's academic knowledge, higher-order reasoning, and problem-solving skills, work attitudes, general employability skills, and the occupational-specific skills necessary for economic independence as a productive and contributing member of society. This term also includes applied technology education. (*Federal Register, #56* (198), Friday, October 11, 1991, Notice of Proposed Rulemaking, § 400.4(b).)

REFERENCES

Congress of the United States Office of Technology Assessment. (1992, February). *Testing in American schools: Asking the right questions* (OTA-SET-520). Washington, DC: U.S. Government Printing Office.

Hill, P., Harvey, J., & Praskac, A. (forthcoming). *Pandora's box: Accountability and performance standards in vocational education*. Berkeley, CA: National Center for Research in Vocational Education, University of California at Berkeley.

Hoachlander, E.G., Kaufman, P., & Levesque, K. (1991). *Vocational education in the United States: 1969–1990*, (92-669). Prepared for the National Center for Education Statistics, U.S. Department of Education.

Hoachlander, E.G., & Rahn, M.L. (1992). *Performance measures and standards for vocational education: 1991*. Berkeley, CA: National Center for Research in Vocational Education, University of California at Berkeley.

McLauglin, M.W. (1991, November). Test-based accountability as a reform strategy. *Phi Delta Kappan, 73*, 248-251.

Mitchell, R. (1992). *Testing for learning: How new approaches to evaluation can improve American schools*. New York, NY: Macmillan/Free Press.

National Center for Research in Vocational Education, University of California at Berkeley. (1992, March). *Workshops on performance measures and standards for vocational education* (video cassette). Berkeley, CA: Author.